



Aggregation-Promoting Factors in Neurodegenerative Diseases

Insights from a *C. elegans* Model

Olga Sin

The research described in this thesis was mostly conducted at the European Research Institute for the Biology of Ageing, University Medical Center Groningen, University of Groningen, The Netherlands.

The research was supported by grants from the European Research Council, the Fundação para a Ciência e Tecnologia (SFRH/BPD/51009/2010), the Groningen University Institute for Drug Exploration (GUIDE) and the Alumni chapter Gooische Groningers facilitated by the Ubbo Emmius Fonds.

The printing of this thesis was financially supported by the University of Groningen, the Graduate School for Drug Exploration (GUIDE), Greiner-Bio-One, Internationale Stichting Alzheimer Onderzoek and Alzheimer Nederland.



Cover concept: Olga Sin and Claudia Arevalo

Cover design and thesis layout: Claudia Arevalo, contact: argo1983@gmail.com
<http://cmgonzalezarevalo.wix.com/layout>

Thesis printed in 100% recycled paper by NetzoDruk, Groningen.

ISBN (print): 78-90-367-8534-1

ISBN (electronic): 978-90-367-8533-4

Copyright (C) 2015 by Olga Sin. All rights reserved. No part of this thesis may be reproduced, stored in a retrieval system, or transmitted in any form or by any means without permission of the author.



rijksuniversiteit
 groningen



umcg



greiner bio-one



alzheimer
 nederland



rijksuniversiteit
 groningen

Aggregation-Promoting Factors in Neurodegenerative Diseases

Insights from a *C. elegans* Model

PhD Thesis

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus Prof. E. Sterken
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Wednesday 13 January 2016 at 12.45 hours

by

Olga Maria Alvarenga Fernandes Sin

born on 27 June 1986
in Macau, China

Supervisor

Prof. dr. E.A.A. Nollen

Co-supervisors

Dr. J. Relvas Bettencourt

Dr. V. Guryev

Assessment Committee

Prof. dr. S. v.d. Heuvel

Prof. dr. G. de Haan

Prof. dr. O. Sibon

Paranymphs:

Manuel Nuno Melo

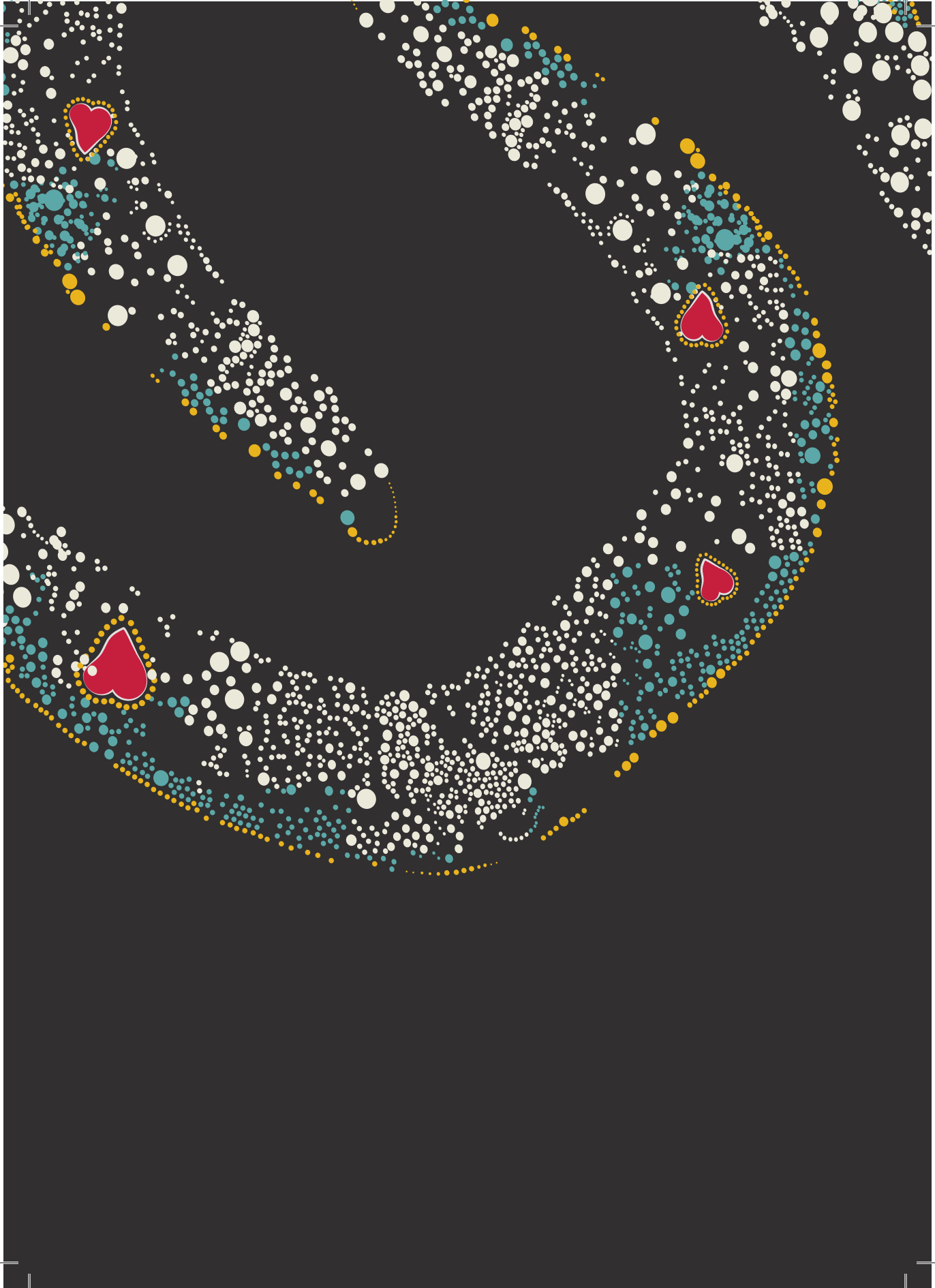
Alejandro Mata Cabana

Table of

Chapter I	Aim and outline of the thesis	9
Chapter II	Regulation of protein homeostasis in neurodegenerative diseases: the role of coding and non-coding genes	13
Chapter III	Genetic screens in <i>Caenorhabditis elegans</i> models for neurodegenerative diseases	55
Chapter IV	Functional sequestration of MOAG-2/LIR-3 by polyglutamine expansion proteins	85
Chapter V	Transcriptional profiling in <i>C. elegans</i> reveals developmental and cellular responses to polyglutamine expansion proteins	115

contents

Chapter VI	Loss of <i>Serf2</i> in mice results in embryonic lethality with incomplete penetrance	133
Chapter VII	Discussion and perspectives	145
Summary		157
Samenvatting		161
Resumo		165
Acknowledgments		170
List of publications		180





CHAPTER I

Aim and outline of the thesis



1. Aim

Neurodegenerative diseases are complex human diseases that are still poorly understood in terms of their etiology and pathogenesis. The objective of this work is to zoom in on the molecular and cellular events that lead to a distinctive characteristic shared by several neurodegenerative diseases: protein aggregation. We take advantage of the nematode *Caenorhabditis elegans* (*C. elegans*) to discover genes that modify protein aggregation and toxicity (proteotoxicity) and to further examine how these genes contribute to the aggregation process. In addition, we aim to validate our observations in the mammalian system, by testing whether the mammalian orthologs of the new genes have similar effects on protein aggregation. Ultimately, we hope that the discovery of these modifiers will add to insights in the field, thereby providing a better understanding of how the aggregation process works, which cellular functions are involved and, importantly, how these are affected in the context of disease.

2. Outline of the thesis

In **Chapter II** we start this thesis by highlighting the importance of a balanced proteome and describing the cellular processes involved in its proper maintenance. We then explain how this balance is disrupted when protein misfolding and aggregation occurs, exemplified by neurodegenerative diseases. Next, we demonstrate the value of using small model organisms including yeast, nematode and fly when identifying genetic modifiers of proteotoxicity and highlight the validation of such modifiers in higher organisms. From a different perspective, we also present an emerging concept, namely the contribution of non-coding RNAs to neurodegeneration, and exemplify how impaired RNA metabolism can lead to neurological disorders.

In **Chapter III**, we focus on *C. elegans* models of neurodegenerative diseases and explain how these can be used to discover modifiers of proteotoxicity in high-throughput genetic screens. We describe the two most frequently applied methods of genetic screening – EMS mutagenesis and RNA interference – and compare these methods by explaining the advantages and disadvantages relative to each other. We conclude by providing examples of genetic screens that enabled the identification of modifiers of proteotoxicity in *C. elegans* models of Parkinson's, Alzheimer's and polyglutamine disease.

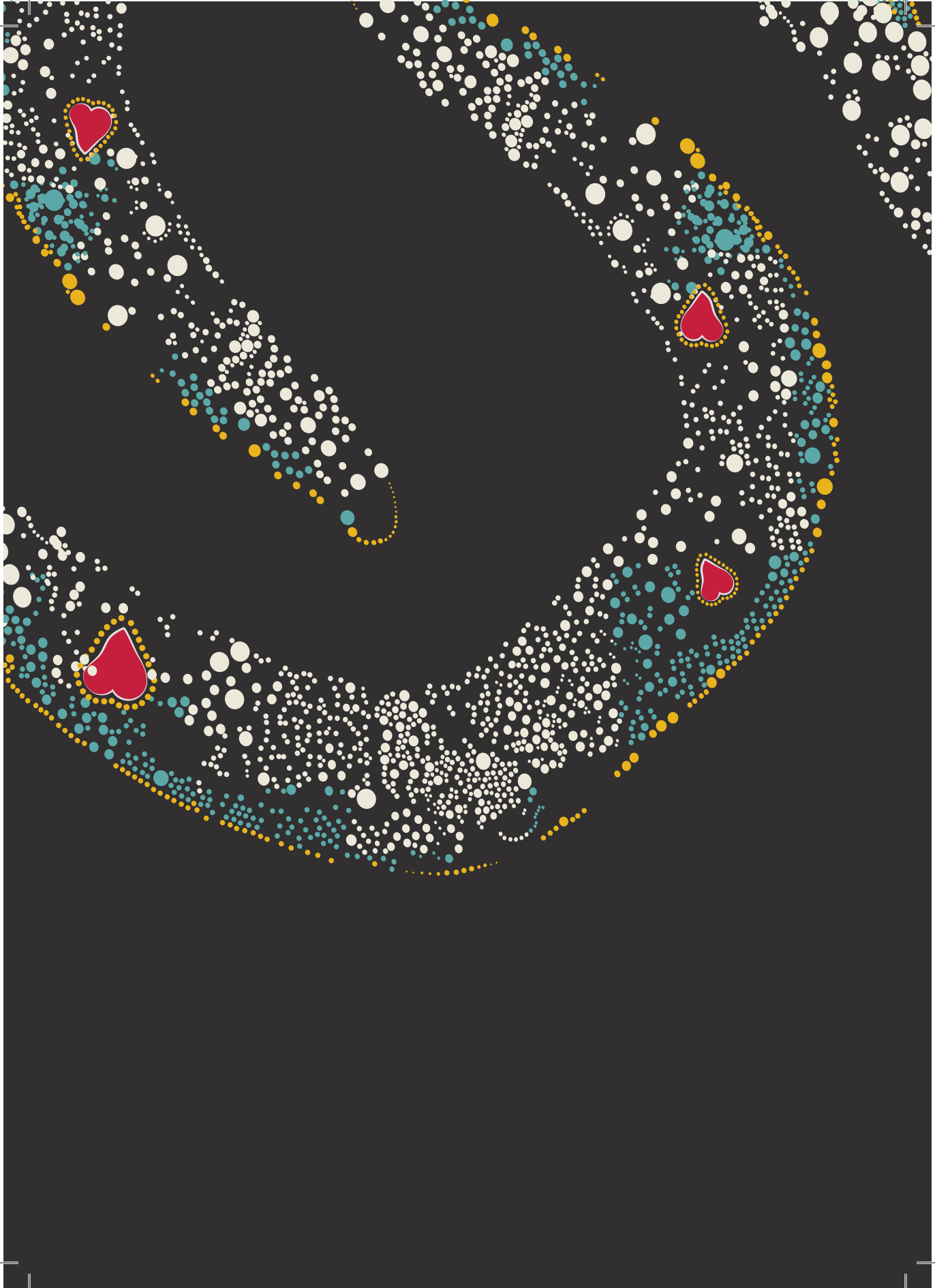
In **Chapter IV**, we describe the characterization of MOAG-2, a modifier of aggregation that was previously identified in a forward genetic screen performed in a *C. elegans* model of polyglutamine diseases. We discovered that MOAG-2 corresponded with LIR-3, a protein of unknown function previously identified in *C. elegans*. We observed that a mutation or a partial deletion in this gene suppressed aggregation in our model. MOAG-2/LIR-3 has two non-canonical C2H2 domains, which are commonly found in transcription factors. This suggested that MOAG-2/LIR-3 might regulate protein aggregation by functioning as a transcription factor. We discovered that MOAG-2/LIR-3 is indeed an RNA Polymerase III-associated transcription factor that regulates the transcription of small non-coding RNAs, including small nucleolar RNAs and transfer RNAs. In this study, we discovered that polyglutamine expansion proteins suppress MOAG-2/LIR-3-mediated transcriptional regulation and convert the MOAG-2/LIR-3 protein into an aggregation-promoting factor.

In **Chapter V**, we focus on how cells respond to proteotoxic stress provoked by the expression of aggregation-prone proteins. To investigate the cellular pathways involved in this stress response, we performed whole transcriptome profiling in

wild type worms and worms expressing polyglutamine expansion proteins. By combining differential gene expression analysis with gene ontology enrichment analysis, we found that the genes that responded to the presence of aggregation-prone proteins in the cell were genes involved in the unfolded protein response, the immune response and oxidative stress. Parallel to this, we also found that aggregation-prone proteins affect *C. elegans* development.

Another modifier of aggregation identified during previous genetic screens was MOAG-4. MOAG-4 has two human orthologs: SERF1 A has 50% amino acid similarity to MOAG-4 and SERF2 has 54%. In **Chapter VI** we follow up on previous studies of SERF2 by Van Ham et al. We aimed to study the effect of SERF2 on amyloid-beta aggregation in the brain of mice and discovered that a full-body *Serf2* knockout resulted in embryonic lethality with incomplete penetrance. Indeed, we observed that the *Serf2* knockout allele was not segregated according to the expected Mendelian ratios, suggesting a crucial function for *Serf2* in mouse development. This observation led us to generate a brain-specific *Serf2* knockout mouse, which is viable and fertile.

In **Chapter VII**, we critically analyze this work and discuss how it has helped to address our initial aims. We further address any questions that remain open and discuss future perspectives that may follow from this work.





CHAPTER II

Regulation of protein homeostasis in neurodegenerative diseases: the role of coding and non-coding genes



Olga Sin^{1,2} & Ellen A. A. Nollen¹

¹ University of Groningen, University Medical Centre Groningen, European Research
Institute for the Biology of Aging, 9700 AD Groningen, The Netherlands

² Graduate Program in Areas of Basic and Applied Biology, Abel Salazar Biomedical
Sciences Institute, University of Porto, 4099-003 Porto, Portugal

Abstract

Protein homeostasis is fundamental for cell function and survival, because proteins are involved in all aspects of cellular function, ranging from cell metabolism and cell division to the cell's response to environmental challenges. Protein homeostasis is tightly regulated by the synthesis, folding, trafficking and clearance of proteins, all of which act in an orchestrated manner to ensure proteome stability. The protein quality control system is enhanced by stress response pathways, which take action whenever the proteome is challenged by environmental or physiological stress. Aging, however, damages the proteome, and such proteome damage is thought to be associated with aging- related diseases. In this review, we discuss the different cellular processes that define the protein quality control system and focus on their role in protein conformational diseases. We highlight the power of using small organisms to model neurodegenerative diseases and how these models can be exploited to discover genetic modulators of protein aggregation and toxicity. We also link findings from small model organisms to the situation in higher organisms and describe how some of the genetic modifiers discovered in organisms such as worms are functionally conserved throughout evolution. Finally, we demonstrate that the non- coding genome also plays a role in maintaining protein homeostasis. In all, this review highlights the importance of protein and RNA homeostasis in neurodegenerative diseases.

Protein homeostasis

Protein folding

Maintaining a healthy proteome is important to ensure cell survival and function. The cell maintains a healthy proteome through a series of complex and tightly regulated surveillance systems (Fig. 1). These systems ensure that each protein is properly folded or assembled in a state that is required for it to perform its function in the cell.

After the synthesis of a nascent polypeptide chain, the protein's amino acid sequence determines whether or not the protein becomes folded, and whether or not chaperone proteins are required for its folding (Fig. 1a, b). Some proteins are thought to exist in a predominantly "unfolded", "disordered" or "intrinsically unstructured" state ([1], also reviewed in [2, 3]). Such proteins are typically involved in transcription, in signaling pathways and in protein networks ([4], also reviewed in [5, 6]). In mammals, about half of all possible proteins are predicted to have long disorganized regions and about 25% are estimated as being intrinsically unstructured [2]. Other proteins have domains within their amino acid sequence that can fold spontaneously, whereas other large, multi-subunit proteins require molecular chaperones to assist in folding to their native state, as shown in *in vitro* studies [7-11].

The molecular chaperones that cooperate in the *de novo* folding or refolding process are subdivided into different classes, which include the Hsp70 system, the small chaperones, the chaperonins and the Hsp90 system [11-14]. In the case of *de novo* synthesis, chaperones protect the nascent polypeptide chain from aberrant contacts with other domains of the same proteins and from aggregation with other proteins (Fig. 1b) ([13, 14], also reviewed in [12, 15]). As a protein is synthesized, it is transiently unfolded and its hydrophobic regions are exposed. Hsp70 is able to recognize these regions and it binds to the protein substrate via its peptide-binding site in an ATP-dependent manner (reviewed in [12, 15, 16]). Hsp70 holds the substrate in an extended conformation, stabilizing it and preventing premature misfolding and aggregation. Next, the substrate can be transferred to another chaperone system, such as the

chaperonins, where folding takes place and a three-dimensional structure is acquired (reviewed in [12, 16, 17]).

When misfolded proteins accumulate, unfolded protein responses can increase the levels of chaperones, which are then able to restore the proteins to their properly folded form (Fig. 1c, d, reviewed in [16], [18-21]). Such an accumulation of misfolded protein is just one of the types of stress that can trigger unfolded protein responses. Unfolded protein responses are mechanisms that are highly conserved from yeast to humans and that are induced upon environmental and physiological stress, such as thermal or oxidative stress (reviewed in [22-24]). In one of these pathways thought to respond to the accumulation misfolded proteins in the cytosol, heat shock factor 1 (HSF-1) acts as a master transcriptional regulator. HSF-1 is activated upon phosphorylation, after which it translocates from the cytosol to the nucleus to bind to so-called heat shock elements, thereby upregulating the transcription of heat shock genes. These genes are then translated into proteins that assist in the refolding of misfolded proteins into functionally active proteins, in preventing unspecific interactions, or in mediating their degradation (Fig. 1d) (reviewed in [19, 22]).

Another strategy used by the cell to restore protein homeostasis is the unfolded protein response that is associated with the endoplasmic reticulum (ER) (Fig. 1e, also reviewed in [18, 25, 26]). The ER is the organelle where proteins enter the secretory pathway to acquire post-translational modifications, after which they are delivered to their corresponding organelle, fixed in the plasma membrane or shuttled outside of the cell to perform their function [27]. If misfolded proteins accumulate, the ER-associated degradation (ERAD) pathway is activated through signal transduction pathways that are mediated by three upstream effectors: inositol-requiring protein 1 (IRE1), activating transcription factor (ATF)-6 and PKR-like endoplasmic reticulum kinase (PERK).

IRE1, ATF-6 and PERK mediate three distinct pathways. Firstly, IRE1 is a transmembrane protein kinase that activates itself by auto-phosphorylation and mediates splicing of Hac1 in yeast and XBP-1 in eukaryotes [28-32]. IRE1 is known to promote the transcription of three groups of genes: stress-responsive genes including molecular chaperones and folding enzymes, genes

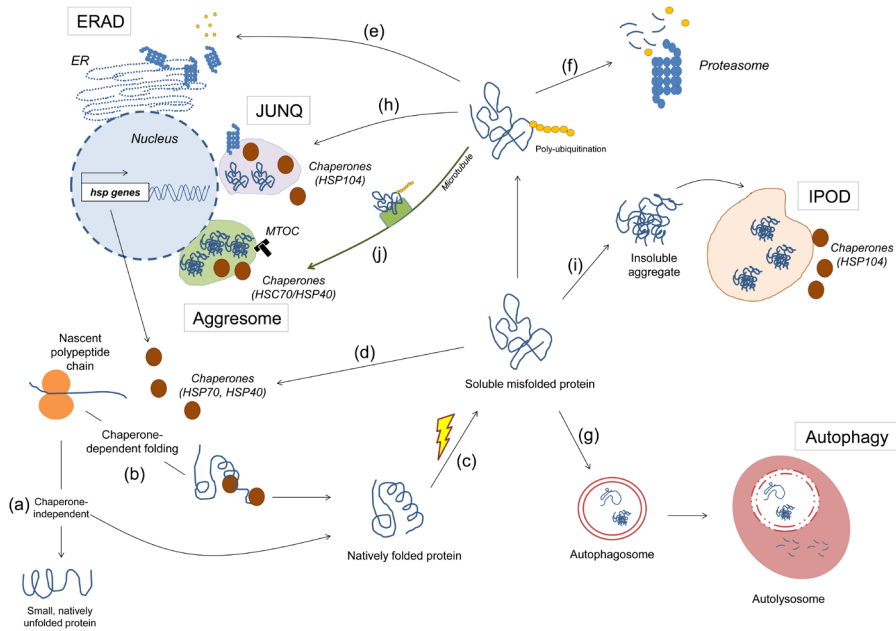


Figure 1. Quality control of cellular proteins. When a protein is synthesized, it can acquire its native state in a chaperone-independent (a) or dependent (b) manner. Upon environmental stress or mutations, the protein may either not acquire its native state or lose it, both leading to misfolding (c). Here, the misfolded protein can be refolded back to its functional conformation with the aid of chaperones (d); or sent to degradation via the ERAD (e), the ubiquitin-proteasome system (f) or autophagy (g). Alternatively, it can be redirected to the JUNQ for posterior refolding or degradation by the proteasome (h) or it can be permanently sequestered in the IPOS (i) or aggregate (j).

involved in ERAD and genes involved in ER trafficking [33-35]. Secondly, ATF-6 is a transmembrane protein with a transcription factor domain (leucine zipper) that translocates from the ER lumen to the Golgi apparatus to be cleaved by proteases [36, 37]. This proteolysis releases the ATF-6 cytosolic fragment, which then enters the nucleus to induce the transcription of ER-resident chaperones and the transcription factor XBP-1, thereby increasing ER protein quality control capacity [29, 37-39]. Thirdly, PERK is a transmembrane kinase protein that phosphorylates the alpha-subunit of the eukaryotic translation initiation factor 2a (eIF2a), thus preventing the binding of the initiator tRNA(Met) to the ribosomal complex, necessary for translation initiation [40-42]. This results in an overall reduction in protein synthesis, thereby attenuating the accumulation of misfolded proteins at the ER.

Protein degradation

If an aberrant protein cannot be folded back into its native state by the molecular chaperones, then it can be eliminated by two proteolytic systems, the proteasome and autophagy (Fig. 1f, g). In the degradation via the ERAD pathway, the ER cooperates tightly with the ubiquitin-proteasome system (UPS) to recognize, mark and traffic the misfolded proteins to the cytosol for degradation (Fig. 1e, reviewed in [18, 43-45]). The exact mechanisms that allow the cell to discriminate misfolded proteins from correctly folded proteins are not fully understood (reviewed in [44, 46, 47]). However, the current notion is that misfolded proteins can be recognized by molecular chaperones (the HSP70 family of proteins) and co-chaperones (the DnaJ/HSP40 family of proteins) [48-51].

An example that illustrates this recognition is the immunoglobulin binding protein (BiP), an HSP70 chaperone that recognizes and binds to the hydrophobic regions of misfolded proteins, thereby preventing their aggregation [49-53]. The binding of the ERAD substrate to BiP and its subsequent release depends on the conversion of ADP to ATP, a process regulated by ERdj proteins, which are part of the DnaJ/Hsp40 family of co-chaperones, and the nucleotide exchange factors GRP170 and BAP/Sil1 [48, 52]. These factors stimulate the ATPase activity of BiP and stabilize its binding to the misfolded protein [54-58]. The ERdj co-chaperones have also been shown to bind directly to unfolded proteins, thus maintaining them in a soluble state to be later recruited by BiP [48, 59]. After the misfolded protein has been identified, it is poly-ubiquitinated to be subsequently targeted for degradation [60-62].

Ubiquitination is a sequential three-step process that marks proteins destined for the proteasome (Fig. 1f). It starts with the activation of ubiquitin (a small 76 amino acid protein) by the activating enzyme E1, followed by binding of ubiquitin to the active site of the ubiquitin-carrier protein E2 and, finally, transfer of the ubiquitin molecule to the substrate in a reaction catalyzed by the ubiquitin protein ligase E3. At least four ubiquitin molecules must be bound to the ERAD substrate for it to be later recognized by the proteasomal machinery [63, 64]. Following this step, the misfolded proteins are delivered to the proteasome (a process called retrotranslocation) and the ubiquitin molecules are removed

from the substrate prior to degradation by the deubiquitinating enzymes and recycled [65-67]. The proteasome is a barrel-shaped, multicatalytic proteinase where proteolysis occurs and proteins are cleaved into peptides 2 to 30 amino acid long [68].

The second proteolytic system, autophagy ("self-eating"), is a cellular degradation mechanism that eliminates cytosolic components, organelles and pathogens via lysosomes (Fig. 1g, [69-72]). It is the part of the cell that ensures protein and organelle turnover, where old cellular components are degraded and recycled molecules become available for cell metabolism [70, 71, 73]. For the purpose of this review, we discuss only the role of autophagy as a protein quality control system.

Autophagy can be classified into three categories: macroautophagy, microautophagy and chaperone-mediated autophagy (CMA). In macroautophagy, a newly formed double membrane vesicle engulfs the cytosolic material, forming the autophagosome. The autophagosome then fuses with an endosome or lysosome, giving rise to the autolysosome where degradation takes place through the action of hydrolytic enzymes (Fig. 1g) [71]. The double membrane that surrounds the autophagosome is derived from the ER, the mitochondria or the plasma membrane [74-78]. In yeast, autophagy is a multi-step process that requires at least 37 autophagy-related (ATG) genes [79-89]. The majority of the ATG genes have shown to be functionally conserved in mammals [90, 91]. In microautophagy, small molecules from the cytoplasm are internalized by the lysosome through invagination of its own membrane [70, 73]. In contrast to autophagy and CMA, much less is known about microautophagy [92].

CMA differs from the former two forms of autophagy in that it does not involve membrane reorganization. Instead, substrates with a KFERQ amino acid motif are recognized by an HSP70 cytosolic chaperone, Hsc70, that binds and delivers them to the CMA receptor at the lysosome [93-96]. Here, the substrate is unfolded before it is translocated into the lumen of the lysosome for degradation, which is assisted by Hsc73, an intralysosomal HSP70 chaperone [97, 98].

Cross talk exists between the UPS and autophagy. Chronic low-level proteasomal inhibition is known to be sufficient to activate autophagy and it has been suggested that ubiquitinated proteins may also be eliminated through this pathway [99-101]. It has also been proposed that macroautophagy may occur as a compensatory mechanism when either the UPS or CMA is impaired [102, 103].

Protein compartmentalization

An alternative pathway for misfolded proteins is the sequestration into specialized protein quality control compartments where they can be either recovered or permanently sequestered (Fig. 1h, i, j) ([104-109], also reviewed in [110, 111]). Distinct quality control compartments harbor different species of misfolded proteins and are evolutionary conserved from yeast to mammals [105, 107-109, 112, 113]. Ubiquitinated misfolded cytosolic proteins are assigned to the juxtanuclear quality control compartment (JUNQ, Fig. 1h). These soluble, mobile misfolded proteins can subsequently be recovered by the molecular chaperone Hsp104 and either refolded back into functionally active proteins or degraded by the proteasomes localized nearby (Fig. 1h) [108, 112]. Non-ubiquitinated misfolded proteins – comprising amyloidogenic proteins – are redistributed to the insoluble protein deposit (IPOD, Fig. 1i). This compartment is localized at the cell periphery and is known to contain insoluble and immobile species, which are not recoverable and seem to remain terminally sequestered there (Fig. 1i) [108]. More recently, it has been proposed that there are no pre-existing compartments in the cell, and that soluble ubiquitinated misfolded proteins (but not the non-ubiquitinated amyloidogenic type) may coalesce and form transient structures termed ‘Q bodies’ that eventually mature into the JUNQ compartments [104].

Much research has focused on finding out whether the redistribution of misfolded proteins to these spatial cytosolic compartments is a random event or whether it depends on the concerted action of sorting factors. Evidence suggests that the latter is the case, and that sorting factors interact with chaperones to deliver misfolded proteins to each compartment [105]. For example, upon physiological stress, Btn2 (a Hook family protein involved in

linking organelles to microtubules) was shown to associate either with the yeast small heat shock protein Hsp42 to assign misfolded proteins to the IPOD or with the chaperone Sis1 to guide misfolded proteins to the JUNQ [105, 107].

Another type of cytosolic compartment – the aggresome – is localized at the microtubule organizing center (MTOC) and is formed when the proteasome is unable to clear misfolded proteins properly (Fig. 1j) [114]. Aggresome formation is accompanied by redistribution of vimentin, an intermediary filament that acquires a cage-like structure in the aggresome. Ubiquitinated misfolded proteins depend on microtubules to be transported to the aggresome, this being done by the dynein/dynactin complex (Fig. 1j) [115]. Interestingly, the JUNQ shares several similarities with the aggresome, including its perinuclear localization, the presence of chaperones and ubiquitinated misfolded proteins [108, 114]. It was also recently shown to functionally associate with the MTOC and vimentin [112]. Indeed, the overwhelming accumulation of misfolded proteins to the JUNQ is thought to render it an immobile, insoluble aggresome over time [112].

Similar structures to aggresomes are the so-called aggresome-like induced structures (ALIS), which were originally discovered in dendritic cells but were later also found in other type of cells [109, 116]. The ALIS is a transient structure with peripheral and juxtanuclear localization. It is induced under a wide variety of stress conditions (e.g. heat shock, starvation, oxidative stress, inflammation) and clusters newly synthesized, ubiquitinated misfolded proteins [106, 109]. ALIS substrates can also be cleared by the proteasome and lysosome [106].

Cell division could be considered as yet another protein quality control system that sequesters misfolded, aggregated proteins (reviewed in [117, 118]). Studies in bacteria and yeast have shown that accumulation of protein aggregates reduces the fitness of these cells, a problem partially resolved by asymmetric division: these protein deposits are retained in the aging mother cell while the daughter cells are freed from damaged proteins, a process also known as replicative rejuvenation [119-123]. In budding yeast, it has been shown that misfolded proteins sorted either to the JUNQ or IPOD remain in the mother

cell after asymmetric cell division, thus avoiding passage of these species onto the daughter cells [124]. Follow-up work from the same group extended this observation to mammalian cells, where the JUNQ (but not the IPOD) continues to be inherited asymmetrically, thereby always freeing one of the two daughter cells from proteotoxicity [112].

While much is now known about the sophisticated quality control mechanisms that the cell has evolved to ensure proper protein homeostasis, several questions remain to be answered. We know that the cell counts on the concerted action of chaperones to avoid an unfolded or misfolded protein interacting aberrantly with other proteins until it can be refolded back into its native state. In case this is not possible, the aberrant protein is sent to be degraded via the ubiquitin-proteasome system or by autophagy. However, it is still not known how the cell chooses one mechanism of degradation over the other or whether the two mechanisms occur simultaneously. Another unknown relates to protein compartmentalization – yet another strategy for putting away proteins that need to be degraded or permanently sequestered. It has not yet been established how the cell can differentiate between degradable and non-degradable proteins and shuttle them to different subcellular compartments. Finally, another important question is how protein quality control changes during aging. Aging itself may be the contributing factor for progressive deterioration of protein homeostasis, impairing the ability of the protein quality control system to handle the equilibrium between protein folding and degradation.

Protein misfolding and aggregation in neurodegenerative diseases

The effects of progressive deterioration of protein homeostasis are thought to play a role in age-related neurodegenerative diseases. The presence of protein aggregates in the brain is namely a hallmark shared by several neurodegenerative diseases, including Parkinson's (PD), Alzheimer's (AD) and Huntington's disease (HD) (reviewed in [125, 126]). In these diseases it is not yet clear why proteins accumulate into aggregates and how this relates to pathogenesis.

Protein aggregation and its relationship to aging and neurodegeneration have

also been widely studied in animal models. Evidence from several animal models suggests that, as the animal ages, the cell's stress response systems become less efficient and less capable of maintaining a balanced proteome [127-133]. This could lead to the progressive accumulation of cytotoxic aggregation-prone disease proteins that cannot be cleared, ultimately resulting in toxicity and cell death [100, 134-137]. In the roundworm *Caenorhabditis elegans*, a model organism much used to study aging, protein aggregation has been shown to occur during aging and to affect the lifespan of the organism [138-140]. As previously discussed, when a protein misfolds it exposes its aggregation-prone domains to the cellular environment – domains that would otherwise be structurally concealed – thereby facilitating the likelihood of aberrant interactions with other proteins, potentially leading to proteotoxicity. Such proteotoxicity is proposed to play a role in protein conformational diseases in humans, including PD, AD and HD.

The type of aggregates that are formed varies for different neurodegenerative diseases. Frontotemporal lobar degeneration with fused in sarcoma is an example of a neurodegenerative disease that is characterized by the presence of amorphous, non-amyloidogenic aggregates ([141, 142], also reviewed in [143]). On the other hand, the common neuropathological feature of PD, AD and HD is the presence of an aggregation-prone disease protein that acquires amyloidogenic properties, causing it to form intracellular amyloid aggregates or extracellular amyloid plaques in the brains of patients (reviewed in [125, 126, 144]). The amyloids present in these neurodegenerative diseases can be distinguished from other amorphous, unstructured aggregates because they are organized, insoluble fibrils with a cross-beta structure and because they can be detected by specific amyloid-binding dyes, namely Congo red and thioflavin T (reviewed in [145, 146]). It is interesting to note that—despite their differences in amino acid sequence and function—several unrelated aggregation-prone disease proteins have one thing in common: in disease they are present as amyloid. This suggests that their ability to form amyloid is related to disease and that they may cause proteotoxicity in a similar manner.

In vitro studies have made clear that virtually any protein can form amyloid fibrils under certain conditions. Such conditions include low pH, high temperature

and high pressure [147-154]. Native proteins are known to exist in equilibrium with their partially unfolded state, and when they are destabilized by certain conditions or mutations, the equilibrium shifts towards amyloid formation. Predicting aggregation-prone regions in proteins is now possible using bioinformatic tools. Examples of such tools are TANGO, which can specifically identify regions prone to form beta sheets, and Waltz, which can distinguish between amyloid sequences and amorphous beta-sheet aggregates [155, 156].

A proposed mechanism for amyloid formation is depicted in Fig. 2. Most of our understanding of this pathway has come from *in vitro* studies of aggregation-prone proteins, including amyloid-beta (seen in AD) and alpha-synuclein (seen in PD) but also from studies of globular proteins, including human lysozyme, superoxide dismutase 1, transthyretin and the acylphosphatase from the archaea *Sulfolobus solfataricus* (reviewed in [125, 146]). One common step of amyloid formation appears to be the conversion of the monomeric, native state protein into an oligomeric intermediate state (Fig. 2). An oligomer is a small and transient cluster of protein molecules that has no fibrillar structure and is of low molecular weight [157-159]. These oligomers can then form protofibrils, which are fibrils 6 to 8 nm in diameter, about 200 nm in length and known to contain beta sheets detectable by Congo red and thioflavin T staining (Fig. 2) [160, 161]. Protofibrils can then convert into amyloid fibrils (Fig. 2) [160]. Of all these aggregation intermediates, it is currently thought that the early ones are cytotoxic and that aggregation may be a neuroprotective response to permanently sequester these intermediates, thereby preventing potentially toxic interactions with other proteins in the cellular milieu [162-165]. In support of this hypothesis, it has been shown that proteins rich in beta-sheet structures aggregate with newly synthesized proteins that have not yet become folded or with intrinsically unfolded proteins, thereby reducing the availability of these proteins to perform their normal function [135]. Further evidence demonstrating that oligomeric or protofibrillar forms of aggregation-prone disease proteins contribute to cell toxicity and death is reviewed elsewhere [144, 146, 166-168].

In a nutshell, the amyloid pathway has only just started to be described and it is not fully understood how protein aggregation correlates with disease. At

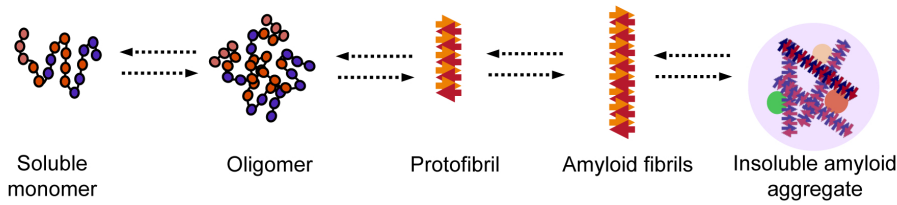


Figure 2. Proposed mechanism for amyloid formation. A protein loses its monomeric native state by conversion into an oligomer which can grow further into amyloidogenic fibrils and ultimately into insoluble amyloid aggregates.

the clinicopathological level, it is striking that there are individuals with high AD pathology (i.e. abundant amyloid deposits and neurofibrillary tangles) and yet do not display any cognitive impairment (reviewed in [169]). This fact makes it difficult to discern what are the boundaries between normal aging and disease pathogenesis. At the cellular and molecular level, what structural properties do aggregation-prone proteins acquire that make them toxic? This question is further complicated by the fact that aggregation-prone proteins such as amyloid-beta, huntingtin or alpha-synuclein do not share sequence, structure or function. A second question is that of how long neuronal cells can deal with these aggregation-prone proteins. And is their slow accumulation in the brain a reflection of an impaired protein quality control system? Finally, the majority of our knowledge about aggregation intermediates has come from *in vitro* studies. It remains to be shown whether oligomeric and fibrillar species exist *in vivo* and what their relevance to pathogenesis is.

Genetic modifiers of proteotoxicity

Genetic screens in small model organisms for protein aggregation in disease

The current understanding of how protein misfolding and aggregation contributes to neurodegeneration is far from complete. Molecular and cellular mechanisms that may regulate neurodegenerative disorders have been discovered in small organisms, the major ones being yeast (*Saccharomyces cerevisiae*), fly (*Drosophila melanogaster*) and nematode (*C. elegans*) (Table 1). In general, these small organisms are easy to grow and manipulate; their genomes are fully sequenced and accessible in public databases; and they provide information relatively quickly due to their short life cycle. Moreover,

the availability of resources such as genome-wide mutant libraries (deletion, overexpression or RNAi-based) further adds to the attraction of using these organisms as powerful genetic tools. Indeed, well-established models of several neurodegenerative diseases, including AD, PD and polyglutamine diseases, have now been generated in each of these small organisms [170]. Of note is that expression of an exogenous aggregation-prone protein typically exclusive to mammals can faithfully mimic some neuropathological features, namely the protein aggregation and toxicity phenotype seen in the diseased brain [170, 171]. And it is this that makes models in small organisms so attractive in the search for evolutionary conserved modifiers of proteotoxicity. These modifiers will provide insight in disease pathology and can be further explored as targets for therapy.

Finding modifiers of proteotoxicity in such models can be relatively quick: researchers can take advantage of high throughput screening techniques using genome-wide overexpression, deletion, or RNAi libraries or using chemical mutagenesis. These resources are unbiased methods that can be used to screen for genes that – when mutated, overexpressed or suppressed – contribute to an increase or decrease of protein aggregation and toxicity. Some of the hits that result from these screens may very well be genes that have already been associated with disease in humans. On the other hand, it is also a way of identifying previously unknown regulators of proteotoxicity – such findings may provide mechanistic insights into that particular disease. It should be noted, however, that genes shown to strongly suppress or enhance aggregation in one model do not always have a similar effect in other models, possibly due to the inherent differences between species or between the methods employed. Nevertheless, functionally conserved genetic modifiers of aggregation and toxicity have been identified across species.

In the end, to establish the value of genes discovered to be involved in aggregation and toxicity in small organisms, the results will have to be reproduced in human neurons and in mammalian animal models. If the function of modifiers of proteotoxicity identified in small organisms is evolutionarily conserved, their mammalian counterparts may become therapeutic targets worthy of future pharmacological investigation

(Table 1). At the same time, small model organisms provide a simple platform that can be used not only to understand the basic mechanisms underlying the causal gene of disease but also as a pharmacological screening tool. Below we describe examples of genetic modifiers that have been studied in different model organisms for PD, AD and polyglutamine diseases.

Parkinson's disease models

Alpha-synuclein is the major constituent of the protein aggregates found in the brains of PD patients, which are also known as Lewy bodies [172]. It is a 140 amino-acid protein that is mostly expressed in the brain and is thought to have a function at the synapse (reviewed in [126, 173]).

The aggregation phenotype is successfully recapitulated in the budding yeast *S. cerevisiae*, where heterologous expression of alpha-synuclein induces toxicity in a concentration-dependent manner and is associated with the formation of cytoplasmic protein aggregates similarly to those observed in the human brain [174]. The characteristics that make yeast a powerful genetic tool for studying neurodegenerative disorders are reviewed elsewhere [171, 175].

In yeast, Cooper et al demonstrated that overexpression and subsequent accumulation of alpha-synuclein impairs vesicle transport from the ER to the Golgi (Table 1) [176]. In the same study, a genome-wide overexpression screen identified the small GTPase Ypt1 as a modifier of alpha-synuclein toxicity. Overexpression of Ypt1p was sufficient to prevent alpha-synuclein toxicity, by enabling forward trafficking from the ER to the Golgi. This observation was further extended to *Drosophila* and *C. elegans* models of PD as well as in rat midbrain primary neurons, where Rab1 – the functionally conserved orthologue of Ypt1p – suppressed dopaminergic neuron loss (Table 1) [176].

Another modifier of proteotoxicity identified from the same original yeast screen was YPK9, an orthologue of the human lysosomal P-type ATPase ATP13A2 (also known as PARK9), an enzyme known to be associated with early onset parkinsonism (Table 1). YPK9 overexpression prevented alpha-synuclein-induced toxicity by reducing intracellular aggregation and restoring alpha-synuclein localization to the plasma membrane [177]. The same study showed

that the *C. elegans* orthologue CATP-6 partially prevented dopaminergic neuron loss, and that knockdown of CATP-6 increased alpha-synuclein misfolding in an age-dependent manner. Finally, in rat primary neuron cultures transduced with a lentivirus carrying the familial alpha-synuclein A53T mutation, heterologous expression of human ATP13A2 prevented neuronal loss (specifically dopaminergic neurons). Notably, this study was the first to show a link between environmental and genetic causes of PD, since YPK9 protected against manganese toxicity in yeast, a heavy metal thought to be risk factor for PD. Indeed, YPK9 was later shown to regulate manganese tolerance via diverse cellular processes, such as vesicle transport, vacuolar organization and chromatin remodeling in yeast (Table 1) [178].

The important role of vesicle-mediated transport in alpha-synuclein toxicity has also been demonstrated by other studies [179, 180]. In a screen performed by Kuwahara et al, the authors discovered ten neuroprotective genes, four of which were involved in endocytosis. Knockdown of two of these genes (*apa-2* and *aps-2*, encoding two different subunits of the AP-2 adaptor protein which mediates clathrin-dependent endocytosis) revealed that deficiencies at synaptic vesicles led to alpha-synuclein neurotoxicity [179].

Several modifiers of proteotoxicity have also been identified using RNAi screens in *C. elegans* models of PD [179-181]. Follow up on this work has revealed *tdo-2* as a general regulator of proteotoxicity and lifespan [182].

Genetic screens not only help us to identify novel modifiers of proteotoxicity, they can also be useful for rediscovering genes that were previously known to be associated with disease. Such an example comes from work by Hamamichi et al, where an RNAi screen identified the autophagy-related gene *Atgr7* as protecting against alpha-synuclein-induced toxicity in *C. elegans* dopaminergic neurons [181]. The mammalian orthologue of *Atgr7* has previously been implicated in neurodegeneration in mice, where it was found to cause axonal degeneration and dystrophy when ablated, thereby highlighting the importance of neuronal autophagy in preventing degeneration (Table 1) [183].

Table 1. Overview of a selection of genetic modifiers of protein aggregation and toxicity that are conserved between small model organisms, mammalian cells and animal models.

Disease	Organism	Transgene	Tissue	Modifier	Mammalian orthologue	Cellular process	Transposed to	Reference
PD	<i>S. cerevisiae</i>	alpha-synuclein	n.a.	YPT1P	RAB1	ER-to-Golgi vesicular trafficking	<i>D.melanogaster</i> , <i>C. elegans</i> and rat midbrain primary neurons	[176]
	<i>S. cerevisiae</i>	alpha-synuclein	n.a.	YPK9	ATP13A2	Manganese homeostasis	<i>C. elegans</i> and rat midbrain primary neurons	[177, 178]
	<i>C. elegans</i>	alpha-synuclein	Body wall muscle and dopaminergic neurons	ATGR7	ATG7	Autophagy	Mouse	[181, 183]
	<i>C. elegans</i>	amyloid-beta	Body wall muscle	DAF-2	IGFR1	Insulin/IGF-1 signaling pathway	Mouse	[195, 197]
AD	<i>S. cerevisiae</i>	amyloid-beta	n.a.	YAP1802	PICALM	Endocytosis	<i>C. elegans</i> , rat cortical neurons, human	[198]
	<i>D.melanogaster</i>	tau	Eye	GLUT1	SLC2A14	Glucose metabolism	Human	[199]
PolyQ	<i>C. elegans</i>	polyglutamine	Body wall muscle	MOAG-4	SERF1A and SERF2	Unknown	Human HEK293 and SK-N-SH cells	[203]
	<i>C. elegans</i>	polyglutamine	Touch receptor neurons	662 modifiers	239 genes, of which 49 differentially expressed in HD mouse brain	Diverse	CHL2 and R6/2 mouse models	[205]
	<i>C. elegans</i>	polyglutamine	Body wall muscle	CCT	TRiC	De novo protein folding	<i>In vitro</i> , <i>S. cerevisiae</i> , mouse N2A neurons, human HeLa cells	[202, 206, 207]
	<i>S. cerevisiae</i>	<i>huntingtin</i> fragment with expanded polyglutamine	n. a.	BNA4	KMO	Kynurenine pathway of tryptophan degradation	<i>D.melanogaster</i> , mouse	[208-210]

Abbreviations: PD - Parkinson's disease; PolyQ - Polyglutamine diseases; AD - Alzheimer's disease; n.a. - not applicable.

Alzheimer's disease models

The brains of patients with AD are characterized by the presence of amyloid plaques and neurofibrillary tangles, which develop as a result of an accumulation of extracellular deposition of two different proteins: amyloid-beta in the plaques and intracellular hyperphosphorylated tau in the tangles (reviewed in [184, 185]). The disease can be caused by a mutation in the gene for amyloid precursor protein (APP), or in presenilin 1 or presenilin 2, all of which alter amyloid production (reviewed in [184, 186]).

C. elegans has been a fundamental tool for dissecting the pathways that link lifespan to AD (Table 1). Specifically, one of the major pathways that regulates lifespan is the insulin/IGF-1 signaling (IIS) pathway – a pathway that has been validated in nematodes, flies and mice and strongly implicated in humans [187–193]. In one of the models that recapitulates AD, *C. elegans* expresses a human amyloid-beta protein fragment (peptide 3–42) in the body wall muscle and progressive paralysis is used as readout for amyloid-beta toxicity [194]. In this model, knockdown of the insulin/IGF-1 receptor DAF-2 not only significantly extended lifespan but also prevented amyloid-beta toxicity by delaying the onset of paralysis, identifying a link between the mechanisms of aging and proteotoxicity [195]. Modulation of lifespan by DAF-2 was also found to be highly dependent on HSF-1 and DAF-16, two transcription factors known to drive the expression of longevity genes [196]. Curiously, while both blocked proteotoxicity, they did so through opposing effects: while HSF-1 promoted disaggregation, DAF-16 pushed aggregation forward, possibly as a means of sequestering the amyloidogenic protein from the cellular milieu [195].

The observation that inhibition of the IIS pathway protects against proteotoxicity was further confirmed in an AD mouse model with haploinsufficiency of IGFR-1, the mammalian orthologue of DAF-2 (Table 1) [197]. Here, reducing only half the expression of IGFR-1 (and thereby the IIS pathway) was sufficient to prevent amyloid-beta toxicity, namely by reducing inflammation and neuron loss. The AD mice with reduced IGFR-1 also performed better in memory and learning tasks than their age-matched AD controls did and this was found to be correlated with the formation of densely packed aggregates in the brain. This

supports the idea that aggregation is a protective mechanism to permanently sequester smaller, soluble oligomeric amyloid-beta species that are proteotoxic.

The importance of modeling neurodegenerative diseases in small organisms has been further reinforced by Treusch et al, who have identified modifiers of amyloid-beta toxicity that are conserved from yeast to humans (Table 1) [198]. Taking advantage of a yeast model of AD, they performed an unbiased genetic screen for modifiers of amyloid-beta toxicity. Of the identified modifiers, six were found to be risk factors for AD in humans – either validated or potential – that had been previously identified from family-based genome-wide association studies (GWAS). These modifiers were specific to amyloid-beta, in that in yeast they did not prevent toxicity induced by another aggregation-prone protein, alpha-synuclein. Another modifier of amyloid-beta toxicity identified by Treusch et al is YAP1802, a suppressor of amyloid-beta proteotoxicity that is involved in clathrin-mediated endocytosis. Its human homolog PICALM is also involved in endocytosis and has been validated as a high risk factor for AD (Table 1). YAP1802 prevents amyloid-beta toxicity in yeast and the human homolog PICALM prevents amyloid-beta toxicity in rat cortical neurons. Notably, this study identifies a causal gene for susceptibility to AD and proposes defective endocytosis as a contributing factor in AD pathology, with a possible role for PICALM.

In another independent study, GWAS data for AD was combined with a functional screen in *Drosophila* (Table 1) [199]. From a set of GWAS variants obtained from patients with AD, Shulman et al found 19 evolutionarily conserved orthologues in the fly that either enhanced or suppressed neurotoxicity associated with tau. Six of these interacted with tau *in vivo*, including the glucose transporter GLUT1, found to be functionally conserved in the human orthologue SLC2A14, further supporting a role for this risk factor as a disease modifying factor (Table 1) [199].

Polyglutamine disease models

In addition to models for PD and AD, there are several other models for aggregation-prone proteins, which include those for human polyglutamine diseases such as Huntington's disease. In polyglutamine diseases trinucleotide

repeats cause expanded tracts of the amino acid glutamine in the encoded protein. In one *C. elegans* model, the animals express expanded glutamine repeats fused to a fluorescent protein in the body wall muscle. Expression of 35 to 40 glutamines is sufficient to cause aggregation, which increases with aging and is correlated with toxicity [200]. This model has been used in at least two genome-wide RNAi screens performed to search for suppressors and enhancers of proteotoxicity [201, 202]. These screens identified genes involved in RNA metabolism, as well as in protein synthesis, folding, trafficking and degradation as polyglutamine modifiers. In a subsequent screen to look for more modifiers, it was found that polyglutamine aggregation is not always coupled with proteotoxicity [201].

In an EMS screen to find genes that drive aggregation, Van Ham et al identified MOAG-4 (modifier of aggregation) as an aggregation-promoting factor in disease models expressing polyglutamine, alpha-synuclein and amyloid-beta, establishing MOAG-4 as a general regulator of proteotoxicity (Table 1) [203]. MOAG-4 is thought to be active during the early steps of the aggregation process, where it drives the formation of compact aggregation intermediates [203]. MOAG-4 is functionally conserved in two human orthologues, SERF1A and SERF2, which have the same aggregation-promoting function in human cell-based models of polyglutamine diseases (Table 1) [203]. Recent insights into the function of one of these proteins, SERF1A, suggest that it acts as an amyloid-promoting factor [204]. In this study, SERF1A recognized a broad range of aggregation-prone proteins (alpha-synuclein, huntingtin, amyloid-beta, prion protein) and mediated their conversion into amyloid *in vitro* [204]. It was further demonstrated that, to do this, SERF1A interacted directly with the monomeric form of the protein to seed amyloid growth, therefore supporting the hypothesis that MOAG-4/SERF1A acts on the early intermediates of the amyloid pathway [204]. SERF1A did not promote aggregation of non-amyloidogenic proteins.

An RNAi screen performed by Lejeune et al identified 662 modifiers that regulate polyglutamine-induced proteotoxicity in *C. elegans* touch receptor neurons, 49 of which were found to be differentially expressed in two mouse

models of HD (Table 1) [205].

Another protein originally identified as a suppressor of polyglutamine aggregation in a *C. elegans* model is the chaperonin CCT [202]. It is composed of eight subunits and, together with HSP70, is involved in *de novo* folding of newly synthesized proteins [12]. Its orthologue, TRiC (also known as TCP), was shown to cooperate with HSP70 to prevent proteotoxicity by promoting the formation of non-toxic, soluble polyglutamine oligomers in a yeast model [206]. TRiC also modulated proteotoxicity in mouse and human cell models (Table 1) [207]. The subunit CCT1 was also shown to physically interact with polyglutamine to suppress aggregation *in vitro*, supporting the hypothesis that TRiC binds to polyglutamine to prevent it from acquiring a potentially toxic conformation [207].

Finally, a modifier identified in yeast is the kynurenine 3-monooxygenase BNA4, whose deletion prevented proteotoxicity induced by mutant huntingtin [208]. Follow-up work showed that genetic ablation or pharmacological inhibition of the orthologue KMO prevented toxicity in a fly and mouse model for HD (Table 1) [209, 210].

In summary, small model organisms including yeast, flies and nematodes are powerful tools for identifying genes involved in protein aggregation and toxicity. Several examples where small animal organisms complement findings from human cell models or mouse models further validate the importance of using these small animal models.

Non-coding RNA in neurodegeneration

When the Human Genome Project started in 1990, it was estimated that 30,000 to 40,000 protein coding genes would be found in the human genome [211]. When the project was completed in 2001, researchers were surprised to find far fewer protein coding genes than expected, namely 21,000, representing only about 2% of the total genome – with the remaining 98% being considered as “junk DNA” [212, 213]. However, it soon became clear that this “junk DNA” actually contained regulatory elements such as non-coding RNA (ncRNA), transcription factor binding sites or certain chromatin structures that govern

gene expression. These conserved functional elements in the human genome were subsequently comprehensively identified and characterized [214]. Within these conserved functional elements, many classes of ncRNA were identified and the list has been growing ever since ([215, 216], also reviewed in [217]). Indeed, the number of ncRNA transcripts are far greater than those coding for proteins and the list of all existing ncRNAs is not yet complete [218]. What we do know is that there are different classes of ncRNA with essential functions in gene transcription, RNA processing and translation, a selection of which is presented in Table 2 (a more complete list can be found in [217]). Indeed, impaired RNA metabolism has been correlated with several neurodegenerative diseases. For example, abnormal expansion repeats in the non-coding regions of disease-related genes induce toxic RNA gain-of-function in myotonic dystrophy, amyotrophic lateral sclerosis and frontotemporal dementia [219, 220]. For the purpose of this review, we focus on a few examples of ncRNAs that have been directly implicated in neurological or neurodegenerative diseases (Table 2).

microRNAs

Over the past few years, it has become evident that ncRNAs are key players in the development and maintenance of the nervous system. Of all classes of ncRNAs identified so far, microRNAs (miRNAs) are those that have been most extensively studied and documented. The function of miRNAs is to bind to the 3'-untranslated region (3' UTR) of messenger RNA and inhibit its translation or target it for degradation (Table 2) [221]. *In situ* hybridization studies in mouse and zebrafish have revealed miRNA to be expressed throughout the brain; these studies have also demonstrated that miRNA expression is spatiotemporally controlled, supporting a biological function for miRNAs in the central nervous system [222-224]. Indeed, several hundreds of miRNAs are involved in brain development [225-229]. miRNAs play a role in virtually every aspect of brain function including neurogenesis, neural differentiation and maintenance, and synaptic plasticity, all of which are described extensively elsewhere [230-232].

Table 2. Classes of RNAs and their localization, size and function.

Class	Localization	Transcribed by	Function	Size	Disease
Messenger RNA (mRNA)	Nucleus and cytoplasm	Pol II	Gene expression into protein	2-5 kb	n.a.
Ribosomal RNA (rRNA)	Cytoplasm	Pol I	Translation, is part of the small or large subunit of the ribosome	70S (prokaryotic) or 80S (eukaryotic)	Unknown
Micro RNA (miRNA)	Nucleus and cytoplasm	Pol II or Pol III	Inhibition of translation or degradation of mRNA	21 – 23 bp	AD, PD, HD
Transfer RNA (tRNA)	Cytoplasm	Pol II or Pol III	Translation, brings cognate amino acid of mRNA to nascent polypeptide chain	70 – 90 bp	PHC-2, PHC-4, neurodevelopmental and neurological disease
Long non-coding RNA (lncRNA)	Nucleus and cytoplasm	Pol II or Pol III	Diverse regulatory roles of other RNAs and proteins; chromatin remodeling	> 200 bp	AD, SCA7
Small nuclear RNA (snRNA)	Nucleus	Pol II or Pol III	Splicing of pre-mRNA	100 - 300 bp	SMA

Abbreviations: AD - Alzheimer's disease; PD - Parkinson's disease; HD - Huntington's disease; PHC - pontocerebellar hypoplasia; SCA - spinocerebellar ataxia; SMA - spinal muscular atrophy; n.a. - not applicable

miRNAs have also been associated with various aspects of aging and neurodegenerative diseases (Table 2) [233– 237]. For example, Northern blotting experiments in the hippocampi of fetuses, adults and AD patients have shown that miRNA expression changes during development and during aging [237]. In these experiments, miR-9 and miR- 128 were upregulated in the AD hippocampus relative to age-matched controls, hinting that these miRNAs may be regulating the expression of genes required for pathogenesis.

At least two other human studies have shown an association between miRNAs and the beta-site APP cleaving enzyme 1 (BACE-1), which is responsible for cleaving APP into the amyloid-beta 1-42 toxic species [235, 236]. In these studies, the expression of miR-107, miR-29a and miR-29b-1 was decreased in the AD brain while expression of BACE-1 was increased (Fig. 4a) [235, 236]. Since these miRNAs target the 30 UTR of BACE-1, it follows that miRNAs can reduce BACE-1 mRNA levels and, therefore, amyloid-beta 1-42 generation in the brain, an effect that is lost in the diseased brain due to the reduced expression of these miRNAs (Fig. 4a) [235, 236]. In a mouse model for AD, miR-34a is thought to

inhibit *bcl-2*, an anti-apoptotic gene that prevents cell death provoked by amyloidogenic species (Fig. 4b) [234]. Additionally, miR-124 has been found to regulate APP alternative splicing in neurons [233].

In PD, downregulation of the miR-34b/c cluster is correlated with downregulation of DJ-1 and Parkin, two genes implicated in the pathogenesis of PD, although a causal link has yet to be determined [238]. It has recently been shown in a cell model that this same cluster directly represses alpha-synuclein mRNA levels and consequently aggregate formation, establishing that miRNAs can have a direct effect on the expression of an aggregation-prone protein [239].

Several miRNAs have also been found to be dysregulated in polyglutamine diseases (Table 2) [240–244]. In HD, REST is a transcription factor that negatively regulates neuronal gene expression and has been found to repress brain-specific miRNAs in mouse and human brains [241, 242]. Two of these miRNAs, miR-9 and miR-9*, have been identified as targeting the REST complex in a negative feedback loop [244]. In a cell model of spinocerebellar ataxia type 1, miR-19, miR-101 and miR-130 cooperatively regulate ataxin-1 expression levels by binding to its 3' UTR [243]. Inhibition of these miRNAs leads to ataxin-1 accumulation in cells and subsequent cell death [243].

tRNAs

Transfer RNAs are essential for mRNA translation into a protein, as they are responsible for transporting the cognate amino acid to the nascent polypeptide chain (Table 2) [245, 246]. Due to the degeneracy of the genetic code, there can be up to five tRNAs per amino acid—termed isoacceptors—that have distinct anticodons for recognizing the same amino acid [245, 246]. On the other hand, tRNAs that share the same anticodon but have distinct body sequences are termed isodecoders, and their number vary greatly [246].

Growing evidence suggests that mutations in individual tRNAs—or in the enzymes involved in their biosynthesis—are a contributing factor in neurodegeneration (Fig. 3) [247–252]. For example, a point mutation (4274T>C) in the mitochondrial tRNA for isoleucine was identified in a patient suffering

from motor neuron disease, although the mechanism by which this mutation might lead to disease is unknown (Fig. 3a) [252]. In a recent study, loss of function of one of the brain-specific tRNA isodecoders for arginine was found to be correlated with neurodegeneration in mice (Fig. 3a) [247]. Specifically, a point mutation (50C>T) in the T loop of the arginine tRNA provoked ribosome stalling, which is normally offset by GTPBP2. However, simultaneous impairment of GTPBP2 in these mice disabled its function as a so-called rescue factor, subsequently resulting in neurodegeneration [247].

Other impairments in the tRNA biosynthesis pathway are seen in pontocerebellar hypoplasia (PHC). PHC is an autosomal recessive neurodegenerative disorder that has six subtypes (PHC1-6) and is generally characterized by hypoplasia and atrophy of the cerebellum and pons [253]. PHC2 and PHC4 arise from impaired tRNA splicing endonuclease (TSEN) activity. TSEN is composed of two catalytic subunits (TSEN 2 and TSEN34) and two non-catalytic subunits (TSEN54 and TSEN15) (Fig. 3c) [250, 254]. It is thought that mutations in both

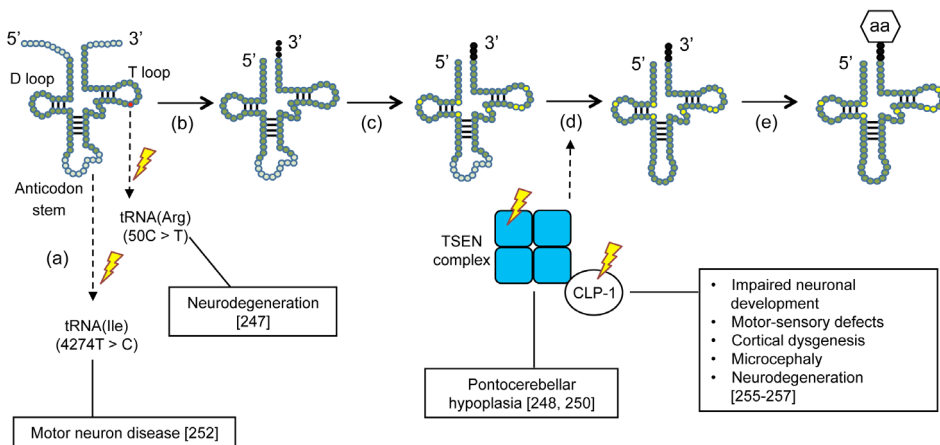


Figure 3. Mutations in the tRNA biosynthesis pathway that lead to neurodegeneration. The point mutation (50C>T) in the T loop of one tRNA isoacceptor for arginine (Arg) provokes neurodegeneration. Another described point mutation (4274T>C) in the mitochondrial tRNA for isoleucine (Ile) has also been associated with motor neuron disease (a). Following transcription, the 5' leader sequence of the pre-tRNA is removed by RNaseP, the 3' end is processed by RNase Z and the trinucleotide CCA is added to the 3' end by a nucleotidyl transferase (b). Different bases of the RNA transcript can undergo chemical modifications (c). The introns of the pre-tRNA are spliced out by a tRNA splicing endonuclease (TSEN). Mutations in these enzymes have been associated with pontocerebellar hypoplasia (PHC) and mutations in their co-factor CLP-1 with motor neuron loss (d). Finally, the mature tRNA is loaded with an amino acid (aa) via tRNA synthetases (e).

catalytic subunits and in TSEN54 may prevent proper complex formation, leading to misplicing of premature tRNAs (pre-tRNAs) into their mature form, thereby unbalancing the tRNA repertoire for protein synthesis [248, 250]. PHC6 results from a mutation in the intronic region of the mitochondrial pre-tRNA synthetase gene for arginine [251].

Finally, CLP-1 is a mammalian kinase that cooperates with the TSEN complex to remove the intronic loop of pre-tRNAs (Fig. 3c) [255]. Loss of CLP-1 results in severe impairment of spinal motor neurons in mice, ultimately leading to respiratory failure [255]. CLP-1 mutations in affected patients have been correlated with neurodevelopment and neurological symptoms in both the central and peripheral nervous system [256, 257].

In summary, these studies demonstrate a crucial role for tRNAs in neuronal function, as either mutations in their transcript or defective post-transcriptional modifications can affect their proper processing and function, ultimately leading to neurodegeneration.

Other ncRNAs

The other non-coding RNAs shown in Table 2 have been less well studied but are nevertheless worthy of mention. Long non-coding RNAs (lncRNAs) are more than 200 nucleotides long and are mostly expressed in the nervous system (Table 2) ([223], also reviewed in [258]). Three lncRNAs have been suggested to be involved in neurodegenerative diseases. Firstly, BACE-1 anti-sense transcript is an lncRNA that competes with miR-485-5p for binding to the BACE-1 mRNA to stabilize it (Fig. 4c) [259]. In AD, the levels of BACE-1 anti-sense transcript are elevated, thereby stabilizing BACE-1 mRNA and enhancing its expression, which further promotes the generation of toxic amyloid-beta 1–42 (Fig. 4c) [259]. Secondly, in spinocerebellar ataxia type 7 (SCA7), lncSCA-7 crosstalks with miR-124 to regulate transcript levels of atxn7 [260]. Thirdly, Abhd11os is an lncRNA that has been shown to be neuroprotective against mutant huntingtin in two mouse models for HD, although the exact mechanism of how this occurs remains to be determined [261].

Small nuclear RNAs (snRNAs) exist as small nuclear ribonucleoproteins (snRNPs) and are major components of the pre-mRNA splicing machinery (Table 2) [262, 263]. The survival motor neuron protein (SMN) is directly involved in the generation of snRNPs [264]. In a mouse model of spinal muscular atrophy, SMN deficiency affects the snRNA pool in a tissue-specific manner, ultimately leading to pre-mRNA splicing defects in a diverse range of genes [264]. Further evidence for the involvement of snRNAs in neurodegeneration comes from work by Jia et al., who revealed that a mutation in a U2 snRNA gene impairs alternative splicing of pre-mRNA which is directly responsible for neuron loss in the cerebellum and hippocampus of mice [265].

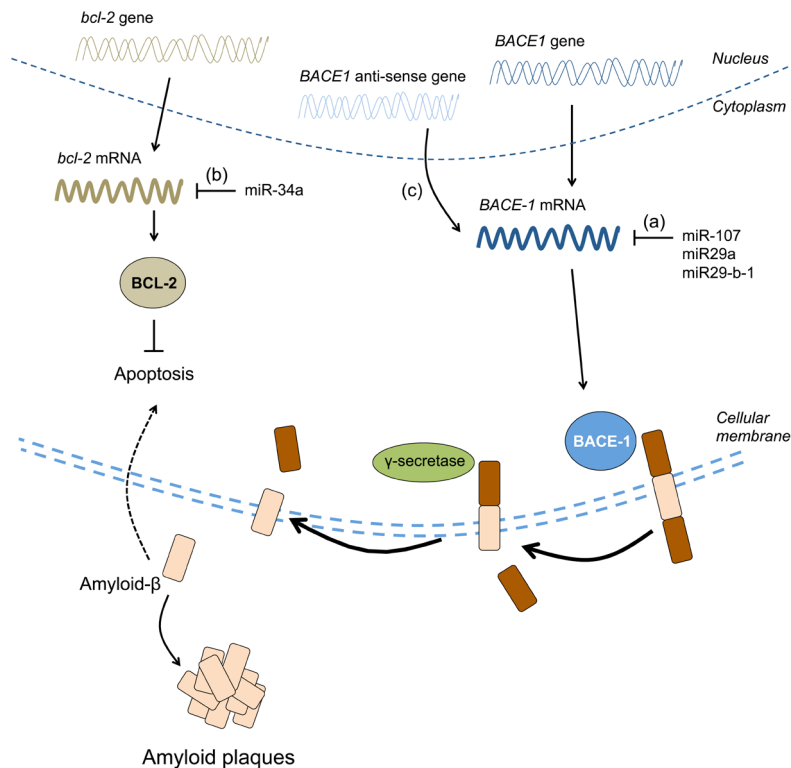


Figure 4. Impaired BACE-1 regulation contributes to AD. miR-107, miR-29a and miR29-b-1 were shown to be decreased in the brain of AD patients while BACE-1 mRNA and protein levels were elevated (a). In an AD mouse model, elevated levels of miR-34a negatively correlate with BCL-2 protein levels, which normally prevent apoptosis induced by amyloid-beta (b). BACE-1 anti-sense transcript was reported to be upregulated in the brain of AD patients. BACE-1 anti-sense transcript stabilizes BACE-1 mRNA thereby facilitating its expression, which ultimately results in the generation of more amyloid-beta (c).

Neurodegeneration is clearly not exclusively caused by imbalances in protein coding genes—it can also arise from dysregulation of ncRNAs. Over the past two decades, we have begun to understand that ncRNAs are not just “transcriptional noise” and have started to define their role in the CNS and in neurodegeneration. Several reports have shown that different classes of ncRNAs influence the expression levels of the disease protein and that each class of ncRNA does so either by affecting the protein post- transcriptionally or through crosstalk with other classes of ncRNAs (miRNAs, lncRNAs). Maintaining a proper environment for protein synthesis is crucial to ensure that each mRNA molecule is effectively spliced and translated into a protein (through tRNAs and snRNAs). To establish the causal relationships between changes in ncRNAs and disease phenotypes, the targets of these ncRNAs need to be uncovered. Understanding the role of ncRNAs will provide insight into the mechanisms of neurodegenerative diseases, which enables the identification of targets for therapeutic interventions.

Acknowledgments

We thank Esther Stroo for critical reading of the manuscript. We thank Sally Hill for editing the manuscript. This work was supported by a European Research Council (ERC) starting Grant (E.A.A.N.) and the Fundação para a Ciência e Tecnologia (SFRH/ BPD/51009/2010) (O.S.).

References

- Weinreb PH, Zhen W, Poon AW, et al. (1996) NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry* 35:13709–13715. doi: 10.1021/bi961799n
- Dunker AK, Silman I, Uversky VN, Sussman JL (2008) Function and structure of inherently disordered proteins. *Curr Opin Struct Biol* 18:756–764. doi: 10.1016/j.sbi.2008.10.002
- Radivojac P, Iakoucheva LM, Oldfield CJ, et al. (2007) Intrinsic disorder and functional proteomics. *Biophysical Journal* 92:1439–1456. doi: 10.1529/biophysj.106.094045
- Iakoucheva LM, Brown CJ, Lawson JD, et al. (2002) Intrinsic Disorder in Cell-signaling and Cancer-associated Proteins. *Journal of Molecular Biology* 323:573–584. doi: 10.1016/S0022-2836(02)00969-5
- Uversky VN, Oldfield CJ, Dunker AK (2005) Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J Mol Recognit* 18:343–384. doi: 10.1002/jmr.747
- Dunker AK, Cortese MS, Romero P, et al. (2005) Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS Journal* 272:5129–5148. doi: 10.1111/j.1742-4658.2005.04948.x
- Agashe VR, Guha S, Chang H-C, et al. (2004) Function of trigger factor and DnaK in multidomain protein folding: increase in yield at the expense of folding speed. *Cell* 117:199–209.
- Frydman J, Erdjument-Bromage H, Tempst P, Hartl FU (1999) Co-translational domain folding as the structural basis for the rapid de novo folding of firefly luciferase. *Nat Struct Biol* 6:697–705. doi: 10.1038/10754
- Freeman BC, Myers MP, Schumacher R, Morimoto RI (1995) Identification of a regulatory motif in Hsp70 that affects ATPase activity, substrate binding and interaction with HDJ-1. *EMBO J* 14:2281–2292.
- Netzer WJ, Hartl FU (1997) Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* 388:343–349. doi: 10.1038/41024
- Schumacher RJ, Hurst R, Sullivan WP, et al. (1994) ATP-dependent chaperoning activity of reticulocyte lysate. *Journal of Biological Chemistry* 269:9493–9499.
- Frydman J (2001) Folding of newly translated proteins in vivo: the role of molecular chaperones. *Annu Rev Biochem* 70:603–647. doi: 10.1146/annurev.biochem.70.1.603
- Kim S, Schilke B, Craig EA, Horwich AL (1998) Folding in vivo of a newly translated yeast cytosolic enzyme is mediated by the SSA class of cytosolic yeast Hsp70 proteins. *Proceedings of the National Academy of Sciences* 95:12860–12865.
- Crombie T, Boyle JP, Coggins JR, Brown AJ (1994) The folding of the bifunctional TRP3 protein in yeast is influenced by a translational pause which lies in a region of structural divergence with *Escherichia coli* indoleglycerol-phosphate synthase. *Eur J Biochem* 226:657–664.
- Hartl FU, Bracher A, Hayer-Hartl M (2011) Molecular chaperones in protein folding and proteostasis. *Nature* 475:324–332. doi: 10.1038/nature10317
- Kim YE, Hipp MS, Bracher A, et al. (2013) Molecular Chaperone Functions in Protein Folding and Proteostasis. *Annu Rev Biochem* 82:323–355. doi: 10.1146/annurev-biochem-060208-092442
- Tyedmers J, Mogk A, Bukau B (2010) Cellular strategies for controlling protein aggregation. *Nature Publishing Group* 11:777–788. doi: 10.1038/nrm2993
- Guerriero CJ, Brodsky JL (2012) The Delicate Balance Between Secreted Protein Folding and Endoplasmic Reticulum-Associated Degradation in Human Physiology. *92:537–576*. doi: 10.1152/physrev.00027.2011
- Westerheide SD, Raynes R, Powell C, et al. (2012) HSF transcription factor family, heat shock response, and protein intrinsic disorder. *Curr Protein Pept Sci* 13:86–103.
- Chen B, Retzlaff M, Roos T, Frydman J (2011) Cellular strategies of protein quality control. *Cold Spring Harb Perspect Biol* 3:a004374–a004374. doi: 10.1101/cshperspect.a004374
- Taylor RC, Dillin A (2011) Aging as an event of proteostasis collapse. *Cold Spring Harb Perspect Biol* 3:a004440–a004440. doi: 10.1101/cshperspect.a004440
- Morimoto RI (2008) Proteotoxic stress and inducible chaperone networks in neurodegenerative disease and aging. *Genes & Development* 22:1427–1438. doi: 10.1101/gad.1657108
- Åkerfelt M, Morimoto RI, Sistonen L (2010) Heat shock factors: integrators of cell stress,

- development and lifespan. 1–11. doi: 10.1038/nrm2938
24. Westerheide SD, Morimoto RI (2005) Heat Shock Response Modulators as Therapeutic Tools for Diseases of Protein Conformation. *Journal of Biological Chemistry* 280:33097–33100. doi: 10.1074/jbc.R500010200
 25. Lin JH, Walter P, Yen TSB (2008) Endoplasmic Reticulum Stress in Disease Pathogenesis. *Annu Rev Pathol Mech Dis* 3:399–425. doi: 10.1146/annurev.pathmechdis.3.121806.151434
 26. Ron D, Walter P (2007) Signal integration in the endoplasmic reticulum unfolded protein response. *Nat Rev Mol Cell Biol* 8:519–529. doi: 10.1038/nrm2199
 27. Lodish H, Berk A, Matsudaira P, et al. (2003) *Molecular Cell Biology*, 5 ed. 1–968.
 28. Calfon M, Zeng H, Urano F, et al. (2002) IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA. *Nature* 415:92–96. doi: 10.1038/415092a
 29. Yoshida H, Matsui T, Yamamoto A, et al. (2001) XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. *Cell* 107:881–891.
 30. Cox JS, Walter P (1996) A novel mechanism for regulating activity of a transcription factor that controls the unfolded protein response. *Cell* 87:391–404.
 31. Cox JS, Shamu CE, Walter P (1993) Transcriptional induction of genes encoding endoplasmic reticulum resident proteins requires a transmembrane protein kinase. *Cell* 73:1197–1206.
 32. Mori K, Ma W, Gething MJ, Sambrook J (1993) A transmembrane protein with a cdc2+/CDC28-related kinase activity is required for signaling from the ER to the nucleus. *Cell* 74:743–756.
 33. Shaffer AL, Shapiro-Shelef M, Iwakoshi NN, et al. (2004) XBP1, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in plasma cell differentiation. *Immunity* 21:81–93. doi: 10.1016/j.immuni.2004.06.010
 34. Okada T, Yoshida H, Akazawa R, et al. (2002) Distinct roles of activating transcription factor 6 (ATF6) and double-stranded RNA-activated protein kinase-like endoplasmic reticulum kinase (PERK) in transcription during the mammalian unfolded protein response. *Biochem J* 366:585–594. doi: 10.1042/BJ20020391
 35. Travers KJ, Patil CK, Wodicka L, et al. (2000) Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. *Cell* 101:249–258.
 36. Ye J, Rawson RB, Komuro R, et al. (2000) ER stress induces cleavage of membrane-bound ATF6 by the same proteases that process SREBPs. *Molecular Cell* 6:1355–1364.
 37. Haze K, Yoshida H, Yanagi H, et al. (1999) Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress. *Mol Biol Cell* 10:3787–3799.
 38. Wu J, Rutkowski DT, Dubois M, et al. (2007) ATF6 α Optimizes Long-Term Endoplasmic Reticulum Function to Protect Cells from Chronic Stress. *Developmental Cell* 13:351–364. doi: 10.1016/j.devcel.2007.07.005
 39. Yoshida H, Okada T, Haze K, et al. (2000) ATF6 activated by proteolysis binds in the presence of NF-Y (CBF) directly to the cis-acting element responsible for the mammalian unfolded protein response. *Molecular and Cellular Biology* 20:6755–6767.
 40. Scheuner D, Song B, McEwen E, et al. (2001) Translational control is required for the unfolded protein response and in vivo glucose homeostasis. *Molecular Cell* 7:1165–1176.
 41. Harding HP, Novoa I, Zhang Y, et al. (2000) Regulated translation initiation controls stress-induced gene expression in mammalian cells. *Molecular Cell* 6:1099–1108.
 42. Harding HP, Zhang Y, Ron D (1999) Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. *Nature* 397:271–274. doi: 10.1038/16729
 43. Brodsky JL (2012) Cleaning up: ER-associated degradation to the rescue. *Cell* 151:1163–1167. doi: 10.1016/j.cell.2012.11.012
 44. Araki K, Nagata K (2011) Protein folding and quality control in the ER. *Cold Spring Harb Perspect Biol* 3:a007526–a007526. doi: 10.1101/cshperspect.a007526
 45. Vembar SS, Brodsky JL (2008) One step at a time: endoplasmic reticulum-associated degradation. *Nat Rev Mol Cell Biol* 9:944–957. doi: 10.1038/nrm2546
 46. Määttänen P, Gehring K, Bergeron JJM, Thomas DY (2010) Protein quality control in the ER: the recognition of misfolded proteins. *Semin Cell Dev Biol* 21:500–511. doi: 10.1016/j.

- semcdb.2010.03.006
47. Goeckeler JL, Brodsky JL (2010) Molecular chaperones and substrate ubiquitination control the efficiency of endoplasmic reticulum-associated degradation. *Diabetes, Obesity and Metabolism* 12:32–38. doi: 10.1111/j.1463-1326.2010.01273.x
 48. Dong M, Bridges JP, Apsley K, et al. (2008) ERdj4 and ERdj5 are required for endoplasmic reticulum-associated protein degradation of misfolded surfactant protein C. *Mol Biol Cell* 19:2620–2630. doi: 10.1091/mbc.E07-07-0674
 49. Nishikawa SI, Fewell SW, Kato Y, et al. (2001) Molecular chaperones in the yeast endoplasmic reticulum maintain the solubility of proteins for retrotranslocation and degradation. *J Cell Biol* 153:1061–1070.
 50. Schmitz A, Maintz M, Kehle T, Herzog V (1995) In vivo iodination of a misfolded proinsulin reveals co-localized signals for Bip binding and for degradation in the ER. *EMBO J* 14:1091–1098.
 51. Knittler MR, Dirks S, Haas IG (1995) Molecular chaperones involved in protein degradation in the endoplasmic reticulum: quantitative interaction of the heat shock cognate protein BiP with partially folded immunoglobulin light chains that are degraded in the endoplasmic reticulum. *Proceedings of the National Academy of Sciences* 92:1764–1768.
 52. Otero JH, Lizák B, Hendershot LM (2010) Life and death of a BiP substrate. *Semin Cell Dev Biol* 21:472–478. doi: 10.1016/j.semcdb.2009.12.008
 53. Blond-Elguindi S, Cwirla SE, Dower WJ, et al. (1993) Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP. *Cell* 75:717–728.
 54. Petrova K, Oyadomari S, Hendershot LM, Ron D (2008) Regulated association of misfolded endoplasmic reticulum luminal proteins with P58/DNAJc3. *EMBO J* 27:2862–2872. doi: 10.1038/emboj.2008.199
 55. Jin Y, Awad W, Petrova K, Hendershot LM (2008) Regulated release of ERdj3 from unfolded proteins by BiP. *EMBO J* 27:2873–2882. doi: 10.1038/emboj.2008.207
 56. Weitzmann A, Volkmer J, Zimmermann R (2006) The nucleotide exchange factor activity of Grp170 may explain the non-lethal phenotype of loss of Sil1 function in man and mouse. *FEBS Lett* 580:5237–5240. doi: 10.1016/j.febslet.2006.08.055
 57. Chung KT, Shen Y, Hendershot LM (2002) BAP, a Mammalian BiP-associated Protein, Is a Nucleotide Exchange Factor That Regulates the ATPase Activity of BiP. *Journal of Biological Chemistry* 277:47557–47563. doi: 10.1074/jbc.M208377200
 58. Liberek K, Marszałek J, Ang D, et al. (1991) *Escherichia coli* DnaJ and GrpE heat shock proteins jointly stimulate ATPase activity of DnaK. *Proceedings of the National Academy of Sciences* 88:2874–2878.
 59. Jin Y, Zhuang M, Hendershot LM (2009) ERdj3, a Luminal ER DnaJ Homologue, Binds Directly to Unfolded Proteins in the Mammalian ER: Identification of Critical Residues. *Biochemistry* 48:41–49. doi: 10.1021/bi8015923
 60. Finley D (2009) Recognition and Processing of Ubiquitin-Protein Conjugates by the Proteasome. *Annu Rev Biochem* 78:477–513. doi: 10.1146/annurev.biochem.78.081507.101607
 61. Hershko A, Ciechanover A (1998) The ubiquitin system. *Annu Rev Biochem* 67:425–479. doi: 10.1146/annurev.biochem.67.1.425
 62. Hochstrasser M (1996) Ubiquitin-dependent protein degradation. *Annu Rev Genet* 30:405–439. doi: 10.1146/annurev.genet.30.1.405
 63. Raasi S, Orlov I, Fleming KG, Pickart CM (2004) Binding of Polyubiquitin Chains to Ubiquitin-associated (UBA) Domains of HHR23A. *Journal of Molecular Biology* 341:1367–1379. doi: 10.1016/j.jmb.2004.06.057
 64. Thrower JS, Hoffman L, Rechsteiner M, Pickart CM (2000) Recognition of the polyubiquitin proteolytic signal. *EMBO J* 19:94–102. doi: 10.1093/emboj/19.1.94
 65. Yao T, Cohen RE (2002) A cryptic protease couples deubiquitination and degradation by the proteasome. *Nature* 419:403–407. doi: 10.1038/nature01071
 66. Verma R, Aravind L, Oania R, et al. (2002) Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. *Science* 298:611–615. doi: 10.1126/science.1075898
 67. Maytal-Kivity V, Reis N, Hofmann K, Glickman MH (2002) MPN+, a putative catalytic motif found in a subset of MPN domain proteins from eukaryotes and prokaryotes, is critical for Rpn11 function. *BMC Biochem* 3:28.
 68. Kisselev AF, Akopian TN, Woo KM, Goldberg AL (1999) The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes. Implications for understanding the degradative mechanism and antigen presentation. *Journal*

- of Biological Chemistry 274:3363–3371. doi: 10.1074/jbc.274.6.3363
69. Mizushima N (2012) Autophagy in Protein and Organelle Turnover. Cold Spring Harbor Symposia on Quantitative Biology 76:397–402. doi: 10.1101/sqb.2011.76.011023
 70. Mizushima N, Komatsu M (2011) Autophagy: Renovation of Cells and Tissues. 147:728–741. doi: 10.1016/j.cell.2011.10.026
 71. Yang Z, Klionsky DJ (2009) An Overview of the Molecular Mechanism of Autophagy. In: Current Topics in Microbiology and Immunology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 1–32
 72. Huang J, Klionsky DJ (2007) Autophagy and human disease. cell cycle 6:1837–1849. doi: 10.4161/cc.6.15.4511
 73. Wong E, Cuervo AM (2010) Integration of clearance mechanisms: the proteasome and autophagy. Cold Spring Harb Perspect Biol 2:a006734–a006734. doi: 10.1101/cshperspect.a006734
 74. Ylä-Anttila P, Vihinen H, Jokitalo E, Eskelinen E-L (2014) 3D tomography reveals connections between the phagophore and endoplasmic reticulum. Autophagy 5:1180–1185. doi: 10.4161/aut.5.8.10274
 75. Hayashi-Nishino M, Fujita N, Noda T, et al. (2009) A subdomain of the endoplasmic reticulum forms a cradle for autophagosome formation. Nat Cell Biol 11:1433–1437. doi: 10.1038/ncb1991
 76. Wang H, Liu J, Zong Y, et al. (2010) miR-106b aberrantly expressed in a double transgenic mouse model for Alzheimer's disease targets TGF- β type II receptor. Brain Research 1357:166–174. doi: 10.1016/j.brainres.2010.08.023
 77. Ravikumar B, Moreau K, Jahreiss L, et al. (2010) Plasma membrane contributes to the formation of pre-autophagosomal structures. Nat Cell Biol 12:747–757. doi: 10.1038/ncb2078
 78. Axe EL, Walker SA, Manifava M, et al. (2008) Autophagosome formation from membrane compartments enriched in phosphatidylinositol 3-phosphate and dynamically connected to the endoplasmic reticulum. J Cell Biol 182:685–701. doi: 10.1083/jcb.200803137
 79. Nazarko TY, Ozeki K, Till A, et al. (2014) Peroxisomal Atg37 binds Atg30 or palmitoyl-CoA to regulate phagophore formation during pexophagy. J Cell Biol 204:541–557. doi: 10.1083/jcb.201307050
 80. Motley AM, Nuttall JM, Hettema EH (2012) Pex3-anchored Atg36 tags peroxisomes for degradation in *Saccharomyces cerevisiae*. EMBO J 31:2852–2868. doi: 10.1038/emboj.2012.151
 81. Nazarko VY, Nazarko TY, Farre J-C, et al. (2011) Atg35, a micropexophagy-specific protein that regulates micropexophagic apparatus formation in *Pichia pastoris*. Autophagy 7:375–385. doi: 10.4161/auto.7.4.14369
 82. Suzuki K, Kondo C, Morimoto M, Ohsumi Y (2010) Selective transport of alpha-mannosidase by autophagic pathways: identification of a novel receptor, Atg34p. J Biol Chem 285:30019–30025. doi: 10.1074/jbc.M110.143511
 83. Kanki T, Wang K, Baba M, et al. (2009) A genomic screen for yeast mutants defective in selective mitochondria autophagy. Mol Biol Cell 20:4730–4738. doi: 10.1091/mbc.E09-03-0225
 84. Kanki T, Wang K, Cao Y, et al. (2009) Atg32 is a mitochondrial protein that confers selectivity during mitophagy. Developmental Cell 17:98–109. doi: 10.1016/j.devcel.2009.06.014
 85. Farre J-C, Manjithaya R, Mathewson RD, Subramani S (2008) PpAtg30 tags peroxisomes for turnover by selective autophagy. Developmental Cell 14:365–376. doi: 10.1016/j.devcel.2007.12.011
 86. Stasyk OV, Stasyk OG, Mathewson RD, et al. (2006) Atg28, a novel coiled-coil protein involved in autophagic degradation of peroxisomes in the methylotrophic yeast *Pichia pastoris*. Autophagy 2:30–38. doi: 10.4161/aut.2226
 87. Kabeya Y, Kawamata T, Suzuki K, Ohsumi Y (2007) Cis1/Atg31 is required for autophagosome formation in *Saccharomyces cerevisiae*. Biochemical and Biophysical Research Communications 356:405–410. doi: 10.1016/j.bbrc.2007.02.150
 88. Kawamata T, Kamada Y, Suzuki K, et al. (2005) Characterization of a novel autophagy-specific gene, ATG29. Biochemical and Biophysical Research Communications 338:1884–1889. doi: 10.1016/j.bbrc.2005.10.163
 89. Klionsky DJ, Cregg JM, Dunn WA, et al. (2003) A unified nomenclature for yeast autophagy-related genes. Developmental Cell 5:539–545.
 90. Yang Z, Klionsky DJ (2010) Eaten alive: a history of macroautophagy. Nat Cell Biol 12:814–822. doi: 10.1038/ncb0910-814
 91. Xie Z, Klionsky DJ (2007) Autophagosome formation: core machinery and adaptations. Nat Cell Biol 9:1102–1109. doi: 10.1038/ncb1007-1102
 92. Mijaljica D, Prescott M, Devenish RJ (2014)

- Microautophagy in mammalian cells: Revisiting a 40-year-old conundrum. *Autophagy* 7:673–682. doi: 10.4161/auto.7.7.14733
93. Koga H, Martinez-Vicente M, Macian F, et al. (2011) A photoconvertible fluorescent reporter to track chaperone-mediated autophagy. *Nat Commun* 2:386. doi: 10.1038/ncomms1393
94. Cuervo AM, Dice JF (1996) A receptor for the selective uptake and degradation of proteins by lysosomes. *Science* 273:501–503.
95. Chiang HL, Terlecky SR, Plant CP, Dice JF (1989) A role for a 70-kilodalton heat shock protein in lysosomal degradation of intracellular proteins. *Science* 246:382–385.
96. Chiang HL, Dice JF (1988) Peptide sequences that target proteins for enhanced degradation during serum withdrawal. *Journal of Biological Chemistry* 263:6797–6805.
97. Agarraberes FA, Terlecky SR, Dice JF (1997) An intralysosomal hsp70 is required for a selective pathway of lysosomal protein degradation. *J Cell Biol* 137:825–834.
98. Salvador N, Aguado C, Horst M, Knecht E (2000) Import of a cytosolic protein into lysosomes by chaperone-mediated autophagy depends on its folding state. *Journal of Biological Chemistry* 275:27447–27456. doi: 10.1074/jbc.M001394200
99. Hara T, Nakamura K, Matsui M, et al. (2006) Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice. *Nature* 441:885–889. doi: 10.1038/nature04724
100. Komatsu M, Waguri S, Chiba T, et al. (2006) Loss of autophagy in the central nervous system causes neurodegeneration in mice. 441:880–884. doi: 10.1038/nature04723
101. Ding Q, Dimayuga E, Martin S, et al. (2004) Characterization of chronic low-level proteasome inhibition on neural homeostasis. *Journal of Neurochemistry* 86:489–497. doi: 10.1046/j.1471-4159.2003.01885.x
102. Pandey UB, Nie Z, Batlevi Y, et al. (2007) HDAC6 rescues neurodegeneration and provides an essential link between autophagy and the UPS. *Nature* 447:859–863. doi: 10.1038/nature05853
103. Massey AC, Kaushik S, Sovak G, et al. (2006) Consequences of the selective blockage of chaperone-mediated autophagy. *Proceedings of the National Academy of Sciences* 103:5805–5810. doi: 10.1073/pnas.0507436103
104. Escusa-Toret S, Vonk WIM, Frydman J (2013) Spatial sequestration of misfolded proteins by a dynamic chaperone pathway enhances cellular fitness during stress. *Nat Cell Biol* 15:1231–1243. doi: 10.1038/ncb2838
105. Malinowska L, Kroschwald S, Munder MC, et al. (2012) Molecular chaperones and stress-inducible protein-sorting factors coordinate the spatiotemporal distribution of protein aggregates. *Mol Biol Cell* 23:3041–3056. doi: 10.1091/mbc.E12-03-0194
106. Liu X-D, Ko S, Xu Y, et al. (2012) Transient aggregation of ubiquitinated proteins is a cytosolic unfolded protein response to inflammation and endoplasmic reticulum stress. *J Biol Chem* 287:19687–19698. doi: 10.1074/jbc.M112.350934
107. Specht S, Miller SBM, Mogk A, Bukau B (2011) Hsp42 is required for sequestration of protein aggregates into deposition sites in *Saccharomyces cerevisiae*. *J Cell Biol* 195:617–629. doi: 10.1083/jcb.201106037
108. Kaganovich D, Kopito R, Frydman J (2008) Misfolded proteins partition between two distinct quality control compartments. *Nature* 454:1088–1095. doi: 10.1038/nature07195
109. Szeto J, Kaniuk NA, Canadien V, et al. (2006) ALIS are stress-induced protein storage compartments for substrates of the proteasome and autophagy. *Autophagy* 2:189–199.
110. Sontag EM, Vonk WI, Frydman J (2014) Sorting out the trash: the spatial nature of eukaryotic protein quality control. *Current Opinion in Cell Biology* 26:139–146. doi: 10.1016/j.ceb.2013.12.006
111. Bagola K, Sommer T (2008) Protein quality control: on IPODs and other JUNQ. *Current Biology* 18:R1019–21. doi: 10.1016/j.cub.2008.09.036
112. Ogrodnik M, Salmonowicz H, Brown R, et al. (2014) Dynamic JUNQ inclusion bodies are asymmetrically inherited in mammalian cell lines through the asymmetric partitioning of vimentin. *Proceedings of the National Academy of Sciences* 111:8049–8054. doi: 10.1073/pnas.1324035111
113. Weisberg SJ, Lyakhovetsky R, Werdiger A-C, et al. (2012) Compartmentalization of superoxide dismutase 1 (SOD1G93A) aggregates determines their toxicity. *Proc Natl Acad Sci USA* 109:15811–15816. doi: 10.1073/pnas.1205829109
114. Johnston JA, Ward CL, Kopito RR (1998) Aggresomes: a cellular response to misfolded proteins. *J Cell Biol* 143:1883–1898.
115. Johnston JA, Illing ME, Kopito RR (2002) Cytoplasmic dynein/dynactin mediates the assembly of aggresomes. *Cell Motil Cytoskeleton* 53:26–38. doi: 10.1002/cm.10057

116. Canadien V, Tan T, Zilber R, et al. (2005) Cutting edge: microbial products elicit formation of dendritic cell aggresome-like induced structures in macrophages. *The Journal of Immunology* 174:2471–2475. doi: 10.4049/jimmunol.174.5.2471
117. Nyström T (2010) Spatial protein quality control and the evolution of lineage-specific ageing. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366:71–75. doi: 10.1098/rstb.2010.0282
118. Sabate R, de Groot NS, Ventura S (2010) Protein folding and aggregation in bacteria. *Cell Mol Life Sci* 67:2695–2715. doi: 10.1007/s00018-010-0344-4
119. Coelho M, Dereli A, Haese A, et al. (2013) Fission Yeast Does Not Age under Favorable Conditions, but Does So after Stress. *Current Biology* 23:1844–1852. doi: 10.1016/j.cub.2013.07.084
120. Liu B, Larsson L, Caballero A, et al. (2010) The Polarisome Is Required for Segregation and Retrograde Transport of Protein Aggregates. *Cell* 140:257–267. doi: 10.1016/j.cell.2009.12.031
121. Rokney A, Shagan M, Kessel M, et al. (2009) E. coli Transports Aggregated Proteins to the Poles by a Specific and Energy-Dependent Process. *Journal of Molecular Biology* 392:589–601. doi: 10.1016/j.jmb.2009.07.009
122. Erjavec N, Cvijovic M, Klipp E, Nyström T (2008) Selective benefits of damage partitioning in unicellular systems and its effects on aging. *Proc Natl Acad Sci USA* 105:18764–18769. doi: 10.1073/pnas.0804550105
123. Lindner AB, Madden R, Demarez A, et al. (2008) Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc Natl Acad Sci USA* 105:3076–3081. doi: 10.1073/pnas.0708931105
124. Spokoini R, Moldavski O, Nahmias Y, et al. (2012) Confinement to Organelle-Associated Inclusion Structures Mediates Asymmetric Inheritance of Aggregated Protein in Budding Yeast. *Cell Rep* 2:738–747. doi: 10.1016/j.celrep.2012.08.024
125. Knowles TPJ, Vendruscolo M, Dobson CM (2014) The amyloid state and its association with protein misfolding diseases. *Nature Publishing Group* 15:384–396. doi: 10.1038/nrm3810
126. Soto C (2003) Unfolding the role of protein misfolding in neurodegenerative diseases. *Nat Rev Neurosci* 4:49–60. doi: 10.1038/nrn1007
127. Andersson V, Hanzén S, Liu B, et al. (2013) Enhancing protein disaggregation restores proteasome activity in aged cells. *Aging (Albany NY)* 5:802–812.
128. Demontis F, Perrimon N (2010) FOXO/4E-BP Signaling in *Drosophila* Muscles Regulates Organism-wide Proteostasis during Aging. *Cell* 143:813–825. doi: 10.1016/j.cell.2010.10.007
129. Ben-Zvi A, Miller EA, Morimoto RI (2009) Collapse of proteostasis represents an early molecular event in *Caenorhabditis elegans* aging. *Proceedings of the National Academy of Sciences* 106:14914–14919. doi: 10.1073/pnas.0902882106
130. Vernace VA, Arnaud L, Schmidt-Glenewinkel T, Figueiredo-Pereira ME (2007) Aging perturbs 26S proteasome assembly in *Drosophila melanogaster*. *The FASEB Journal* 21:2672–2682. doi: 10.1096/fj.06-6751com
131. Ferrington DA, Husom AD, Thompson LV (2005) Altered proteasome structure, function, and oxidation in aged muscle. 19:644–646. doi: 10.1096/fj.04-2578fje
132. Cuervo AM, Dice JF (2000) Age-related Decline in Chaperone-mediated Autophagy. *Journal of Biological Chemistry* 275:31505–31513. doi: 10.1074/jbc.M002102200
133. Vittorini S, Paradiso C, Donati A, et al. (1999) The age-related accumulation of protein carbonyl in rat liver correlates with the age-related decline in liver proteolytic activities. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 54:B318–23.
134. Gupta R, Kasturi P, Bracher A, et al. (2011) Firefly luciferase mutants as sensors of proteome stress. *Nature Methods* 8:879–884. doi: 10.1038/nmeth.1697
135. Olzscha H, Schermann SM, Woerner AC, et al. (2011) Amyloid-like Aggregates Sequester Numerous Metastable Proteins with Essential Cellular Functions. *Cell* 144:67–78. doi: 10.1016/j.cell.2010.11.050
136. Gidalevitz T, Ben-Zvi A, Ho KH, et al. (2006) Progressive disruption of cellular protein folding in models of polyglutamine diseases. *Science* 311:1471–1474. doi: 10.1126/science.1124514
137. Bence NF (2001) Impairment of the Ubiquitin-Proteasome System by Protein Aggregation. *Science* 292:1552–1555. doi: 10.1126/science.292.5521.1552
138. Walther DM, Kasturi P, Zheng M, et al. (2015) Widespread Proteome Remodeling and Aggregation in Aging *C. elegans*. *Cell* 161:919–932. doi: 10.1016/j.cell.2015.03.032

139. Reis-Rodrigues P, Czerwieniec G, Peters TW, et al. (2012) Proteomic analysis of age-dependent changes in protein solubility identifies genes that modulate lifespan. *Aging Cell* 11:120–127. doi: 10.1111/j.1474-9726.2011.00765.x
140. David DC, Ollikainen N, Trinidad JC, et al. (2010) Widespread protein aggregation as an inherent part of aging in *C. elegans*. *PLoS Biol* 8:e1000450. doi: 10.1371/journal.pbio.1000450
141. Bigio EH, Wu JY, Deng H-X, et al. (2013) Inclusions in frontotemporal lobar degeneration with TDP-43 proteinopathy (FTLD-TDP) and amyotrophic lateral sclerosis (ALS), but not FTLD with FUS proteinopathy (FTLD-FUS), have properties of amyloid. *Acta Neuropathol* 125:463–465. doi: 10.1007/s00401-013-1089-6
142. Lashley T, Rohrer JD, Bandopadhyay R, et al. (2011) A comparative clinical, pathological, biochemical and genetic study of fused in sarcoma proteinopathies. *Brain* 134:2548–2564. doi: 10.1093/brain/awr160
143. Shelkvnikova TA (2013) Modelling FUSopathies: focus on protein aggregation. *Biochem Soc Trans* 41:1613–1617. doi: 10.1042/BST20130212
144. Haass C, Selkoe DJ (2007) Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β -peptide. *Nat Rev Mol Cell Biol* 8:101–112. doi: 10.1038/nrm2101
145. Eichner T, Radford SE (2011) A Diversity of Assembly Mechanisms of a Generic Amyloid Fold. 43:8–18. doi: 10.1016/j.molcel.2011.05.012
146. Chiti F, Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* 75:333–366. doi: 10.1146/annurev.biochem.75.101304.123901
147. De Felice FG (2004) Formation of amyloid aggregates from human lysozyme and its disease-associated variants using hydrostatic pressure. *The FASEB Journal*. doi: 10.1096/fj.03-1072fje
148. Schmittschmitt JP, Scholtz JM (2003) The role of protein stability, solubility, and net charge in amyloid fibril formation. *Protein Sci* 12:2374–2378. doi: 10.1110/ps.03152903
149. Fändrich M, Fletcher MA, Dobson CM (2001) Amyloid fibrils from muscle myoglobin. *Nature* 410:165–166. doi: 10.1038/35065514
150. Ferrão-Gonzales AD, Souto SO, Silva JL, Foguel D (2000) The preaggregated state of an amyloidogenic protein: hydrostatic pressure converts native transthyretin into the amyloidogenic state. *Proceedings of the National Academy of Sciences* 97:6445–6450.
151. McParland VJ, Kad NM, Kalverda AP, et al. (2000) Partially Unfolded States of β 2-Microglobulin and Amyloid Formation in Vitro. *Biochemistry* 39:8735–8746. doi: 10.1021/bi000276j
152. Chiti F, Webster P, Taddei N, et al. (1999) Designing conditions for in vitro formation of amyloid protofilaments and fibrils. *Proceedings of the National Academy of Sciences* 96:3590–3594.
153. Guijarro JI, Sunde M, Jones JA, et al. (1998) Amyloid fibril formation by an SH3 domain. *Proceedings of the National Academy of Sciences* 95:4224–4228.
154. Litvinovich SV, Brew SA, Aota S, et al. (1998) Formation of amyloid-like fibrils by self-association of a partially unfolded fibronectin type III module. *Journal of Molecular Biology* 280:245–258. doi: 10.1006/jmbi.1998.1863
155. Maurer-Stroh S, Debulpaep M, Kuemmerer N, et al. (2010) Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. *Nature Methods* 7:237–242. doi: 10.1038/nmeth.1432
156. Fernandez-Escamilla A-M, Rousseau F, Schymkowitz J, Serrano L (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat Biotechnol* 22:1302–1306. doi: 10.1038/nbt1012
157. Walsh DM, Klyubin I, Fadeeva JV, et al. (2002) Naturally secreted oligomers of amyloid beta protein potently inhibit hippocampal long-term potentiation in vivo. *Nature* 416:535–539. doi: 10.1038/416535a
158. Lambert MP, Barlow AK, Chromy BA, et al. (1998) Diffusible, nonfibrillar ligands derived from A β 1-42 are potent central nervous system neurotoxins. *Proceedings of the National Academy of Sciences* 95:6448–6453.
159. Kuo YM, Emmerling MR, Vigo-Pelfrey C, et al. (1996) Water-soluble A β (N-40, N-42) oligomers in normal and Alzheimer disease brains. *Journal of Biological Chemistry* 271:4077–4081.
160. Walsh DM, Hartley DM, Kusumoto Y, et al. (1999) Amyloid beta-protein fibrillogenesis. Structure and biological activity of protofibrillar intermediates. *Journal of Biological Chemistry* 274:25945–25952. doi: 10.1074/jbc.274.36.25945
161. Walsh DM, Lomakin A, Benedek GB, et al. (1997) Amyloid beta-protein fibrillogenesis. Detection of a protofibrillar intermediate. *Journal of*

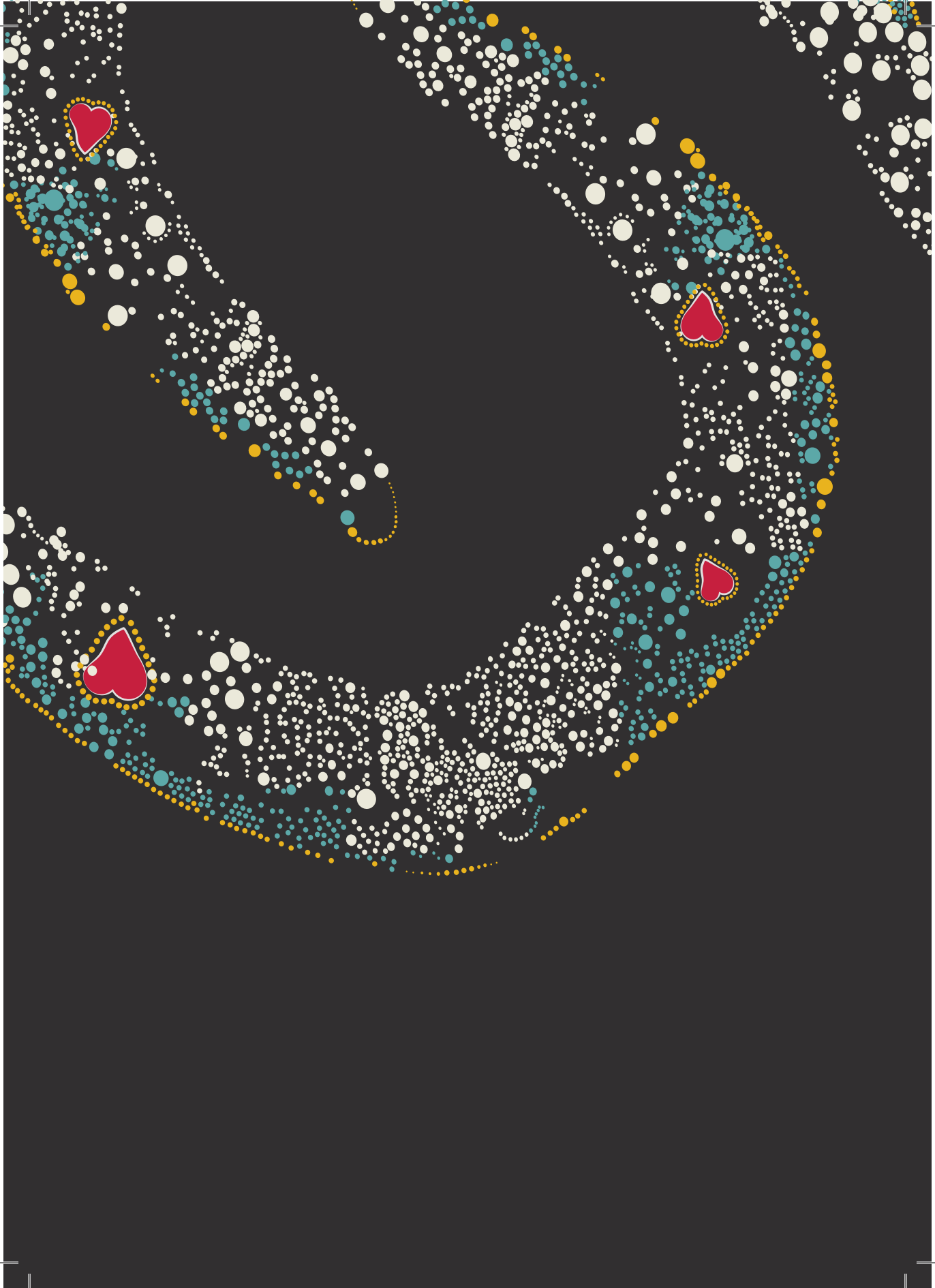
- Biological Chemistry 272:22364–22372.
162. Miller J, Arrasate M, Brooks E, et al. (2011) Identifying polyglutamine protein species in situ that best predict neurodegeneration. *Nat Chem Biol* 7:925–934. doi: 10.1038/nchembio.694
 163. Bolognesi B, Kumita JR, Barros TP, et al. (2010) ANS Binding Reveals Common Features of Cytotoxic Amyloid Species. 5:735–740. doi: 10.1021/cb1001203
 164. Arrasate M, Mitra S, Schweitzer ES, et al. (2004) Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* 431:805–810. doi: 10.1038/nature02998
 165. Kaye R, Head E, Thompson JL, et al. (2003) Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* 300:486–489. doi: 10.1126/science.1079469
 166. De Genst E, Messer A, Dobson CM (2014) Antibodies and protein misfolding: From structural research tools to therapeutic strategies. *BBA - Proteins and Proteomics* 1844:1907–1919. doi: 10.1016/j.bbapap.2014.08.016
 167. Benilova I, Karran E, De Strooper B (2012) The toxic A β oligomer and Alzheimer's disease: an emperor in need of clothes. *Nat Neurosci* 15:349–357. doi: 10.1038/nn.3028
 168. Lashuel HA, Lansbury PT (2006) Are amyloid diseases caused by protein aggregates that mimic bacterial pore-forming toxins? *Q Rev Biophys* 39:167–201. doi: 10.1017/S0033583506004422
 169. Serrano-Pozo A, Frosch MP, Masliah E, Hyman BT (2011) Neuropathological alterations in Alzheimer disease. *Cold Spring Harb Perspect Med* 1:a006189. doi: 10.1101/cshperspect.a006189
 170. van Ham TJ, Breitling R, Swertz MA, Nollen EAA (2009) Neurodegenerative diseases: Lessons from genome-wide screens in small model organisms. *EMBO Mol Med* 1:360–370. doi: 10.1002/emmm.200900051
 171. Miller-Fleming L, Giorgini F, Outeiro TF (2008) Yeast as a model for studying human neurodegenerative disorders. *Biotechnol J* 3:325–338. doi: 10.1002/biot.200700217
 172. Spillantini MG, Schmidt ML, Lee VM, et al. (1997) Alpha-synuclein in Lewy bodies. *Nature* 388:839–840. doi: 10.1038/42166
 173. Kim WS, Kågedal K, Halliday GM (2014) Alpha-synuclein biology in Lewy body diseases. *Alzheimers Res Ther* 6:73. doi: 10.1186/s13195-014-0073-2
 174. Outeiro TF, Lindquist S (2003) Yeast cells provide insight into alpha-synuclein biology and pathobiology. *Science* 302:1772–1775. doi: 10.1126/science.1090439
 175. Winderickx J, Delay C, De Vos A, et al. (2008) Protein folding diseases and neurodegeneration: lessons learned from yeast. *Biochim Biophys Acta* 1783:1381–1395. doi: 10.1016/j.bbamcr.2008.01.020
 176. Cooper AA, Gitler AD, Cashikar A, et al. (2006) Alpha-synuclein blocks ER-Golgi traffic and Rab1 rescues neuron loss in Parkinson's models. *Science* 313:324–328. doi: 10.1126/science.1129462
 177. Gitler AD, Chesi A, Geddie ML, et al. (2009) Alpha-Synuclein is part of a diverse and highly conserved interaction network that includes PARK9 and manganese toxicity. *Nat Genet* 41:308–315. doi: 10.1038/ng.300
 178. Chesi A, Kilaru A, Fang X, et al. (2012) The role of the Parkinson's disease gene PARK9 in essential cellular pathways and the manganese homeostasis network in yeast. *PLoS ONE* 7:e34178. doi: 10.1371/journal.pone.0034178
 179. Kuwahara T, Koyama A, Koyama S, et al. (2008) A systematic RNAi screen reveals involvement of endocytic pathway in neuronal dysfunction in alpha-synuclein transgenic *C. elegans*. *Human Molecular Genetics* 17:2997–3009. doi: 10.1093/hmg/ddn198
 180. van Ham TJ, Thijssen KL, Breitling R, et al. (2008) *C. elegans* model identifies genetic modifiers of alpha-synuclein inclusion formation during aging. *PLoS Genet* 4:e1000027. doi: 10.1371/journal.pgen.1000027
 181. Hamamichi S, Rivas RN, Knight AL, et al. (2008) Hypothesis-based RNAi screening identifies neuroprotective genes in a Parkinson's disease model. *Proc Natl Acad Sci USA* 105:728–733. doi: 10.1073/pnas.0711018105
 182. van der Goot AT, Zhu W, Vazquez-Manrique RP, et al. (2012) Delaying aging and the aging-associated decline in protein homeostasis by inhibition of tryptophan degradation. *Proc Natl Acad Sci USA* 109:14912–14917. doi: 10.1073/pnas.1203083109
 183. Komatsu M, Wang QJ, Holstein GR, et al. (2007) Essential role for autophagy protein Atg7 in the maintenance of axonal homeostasis and the prevention of axonal degeneration. *Proceedings of the National Academy of Sciences* 104:14489–14494. doi: 10.1073/pnas.0701311104

184. Ballard C, Gauthier S, Corbett A, et al. (2011) Alzheimer's disease. *Lancet* 377:1019–1031. doi: 10.1016/S0140-6736(10)61349-9
185. Takashima A (2009) Amyloid-beta, tau, and dementia. *J Alzheimers Dis* 17:729–736. doi: 10.3233/JAD-2009-1090
186. Bettens K, Sleegers K, Van Broeckhoven C (2013) Genetic insights in Alzheimer's disease. *The Lancet Neurology* 12:92–104. doi: 10.1016/S1474-4422(12)70259-4
187. Flachsbarf F, Caliebe A, Kleindorp R, et al. (2009) Association of FOXO3A variation with human longevity confirmed in German centenarians. *Proc Natl Acad Sci USA* 106:2700–2705. doi: 10.1073/pnas.0809594106
188. Suh Y, Atzmon G, Cho M-O, et al. (2008) Functionally significant insulin-like growth factor I receptor mutations in centenarians. *Proceedings of the National Academy of Sciences* 105:3438–3442. doi: 10.1073/pnas.0705467105
189. Willcox BJ, Donlon TA, He Q, et al. (2008) FOXO3A genotype is strongly associated with human longevity. *Proc Natl Acad Sci USA* 105:13987–13992. doi: 10.1073/pnas.0801030105
190. Blüher M, Kahn BB, Kahn CR (2003) Extended longevity in mice lacking the insulin receptor in adipose tissue. *Science* 299:572–574. doi: 10.1126/science.1078223
191. Holzenberger M, Dupont J, Ducos B, et al. (2003) IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. *Nature* 421:182–187. doi: 10.1038/nature01298
192. Tatar M (2001) A Mutant *Drosophila* Insulin Receptor Homolog That Extends Life-Span and Impairs Neuroendocrine Function. *Science* 292:107–110. doi: 10.1126/science.1057987
193. Kenyon C, Chang J, Gensch E, et al. (1993) A *C. elegans* mutant that lives twice as long as wild type. *Nature* 366:461–464. doi: 10.1038/366461a0
194. Link CD (1995) Expression of human beta-amyloid peptide in transgenic *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences* 92:9368–9372.
195. Cohen E, Bieschke J, Perciavalle RM, et al. (2006) Opposing activities protect against age-onset proteotoxicity. *Science* 313:1604–1610. doi: 10.1126/science.1124646
196. Hsu A-L, Murphy CT, Kenyon C (2003) Regulation of aging and age-related disease by DAF-16 and heat-shock factor. *Science* 300:1142–1145. doi: 10.1126/science.1083701
197. Cohen E, Paulsson JF, Blinder P, et al. (2009) Reduced IGF-1 signaling delays age-associated proteotoxicity in mice. *Cell* 139:1157–1169. doi: 10.1016/j.cell.2009.11.014
198. Treusch S, Hamamichi S, Goodman JL, et al. (2011) Functional links between A β toxicity, endocytic trafficking, and Alzheimer's disease risk factors in yeast. *Science* 334:1241–1245. doi: 10.1126/science.1213210
199. Shulman JM, Chipendo P, Chibnik LB, et al. (2011) Functional screening of Alzheimer pathology genome-wide association signals in *Drosophila*. *Am J Hum Genet* 88:232–238. doi: 10.1016/j.ajhg.2011.01.006
200. Morley JF, Brignull HR, Weyers JJ, Morimoto RI (2002) The threshold for polyglutamine-expansion protein aggregation and cellular toxicity is dynamic and influenced by aging in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences* 99:10417–10422. doi: 10.1073/pnas.152161099
201. Silva MC, Fox S, Beam M, et al. (2011) A genetic screening strategy identifies novel regulators of the proteostasis network. *PLoS Genet* 7:e1002438. doi: 10.1371/journal.pgen.1002438
202. Nollen EAA, Garcia SM, van Haaften G, et al. (2004) Genome-wide RNA interference screen identifies previously undescribed regulators of polyglutamine aggregation. *Proceedings of the National Academy of Sciences* 101:6403–6408. doi: 10.1073/pnas.0307697101
203. van Ham TJ, Holmberg MA, van der Goot AT, et al. (2010) Identification of MOAG-4/SERF as a regulator of age-related proteotoxicity. *Cell* 142:601–612. doi: 10.1016/j.cell.2010.07.020
204. Falsone SF, Meyer NH, Schrank E, et al. (2012) SERF protein is a direct modifier of amyloid fiber assembly. *Cell Rep* 2:358–371. doi: 10.1016/j.celrep.2012.06.012
205. Lejeune F-X, Mesrob L, Parmentier F, et al. (2012) Large-scale functional RNAi screen in *C. elegans* identifies genes that regulate the dysfunction of mutant polyglutamine neurons. *BMC Genomics* 13:91. doi: 10.1186/1471-2164-13-91
206. Behrends C, Langer CA, Boteva R, et al. (2006) Chaperonin TRiC promotes the assembly of polyQ expansion proteins into nontoxic oligomers. *Molecular Cell* 23:887–897. doi: 10.1016/j.molcel.2006.08.017
207. Tam S, Geller R, Spiess C, Frydman J (2006) The chaperonin TRiC controls polyglutamine aggregation and toxicity through subunit-

- specific interactions. *Nat Cell Biol* 8:1155–1162. doi: 10.1038/ncb1477
208. Giorgini F, Guidetti P, Nguyen Q, et al. (2005) A genomic screen in yeast implicates kynurenine 3-monooxygenase as a therapeutic target for Huntington disease. *Nat Genet* 37:526–531. doi: 10.1038/ng1542
 209. Campesan S, Green EW, Breda C, et al. (2011) The kynurenine pathway modulates neurodegeneration in a *Drosophila* model of Huntington's disease. *Curr Biol* 21:961–966. doi: 10.1016/j.cub.2011.04.028
 210. Zwilling D, Huang S-Y, Sathyaikumar KV, et al. (2011) Kynurenine 3-monooxygenase inhibition in blood ameliorates neurodegeneration. *Cell* 145:863–874. doi: 10.1016/j.cell.2011.05.020
 211. Lander ES, Linton LM, Birren B, et al. (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860–921. doi: 10.1038/35057062
 212. Clamp M, Fry B, Kamal M, et al. (2007) Distinguishing protein-coding and noncoding genes in the human genome. *Proc Natl Acad Sci USA* 104:19428–19433. doi: 10.1073/pnas.0709013104
 213. Mouse Genome Sequencing Consortium, Waterston RH, Lindblad-Toh K, et al. (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* 420:520–562. doi: 10.1038/nature01262
 214. ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489:57–74. doi: 10.1038/nature11247
 215. Kapranov P, Cheng J, Dike S, et al. (2007) RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription. *Science* 316:1484–1488. doi: 10.1126/science.1138341
 216. Carninci P, Kasukawa T, Katayama S, et al. (2005) The transcriptional landscape of the mammalian genome. *Science* 309:1559–1563. doi: 10.1126/science.1112014
 217. Cech TR, Steitz JA (2014) The Noncoding RNA Revolution—Trashing Old Rules to Forge New Ones. *Cell* 157:77–94. doi: 10.1016/j.cell.2014.03.008
 218. Kapranov P, Laurent GS, Raz T, et al. (2010) The majority of total nuclear-encoded non-ribosomal RNA in a human cell is “dark matter” un-annotated RNA. *BMC Biology* 8:149. doi: 10.1186/1741-7007-8-149
 219. Donnelly CJ, Zhang P-W, Pham JT, et al. (2013) RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. *Neuron* 80:415–428. doi: 10.1016/j.neuron.2013.10.015
 220. Miller JW, Urbinati CR, Teng-Umuay P, et al. (2000) Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J* 19:4439–4448. doi: 10.1093/emboj/19.17.4439
 221. Krol J, Loedige I, Filipowicz W (2010) The widespread regulation of microRNA biogenesis, function and decay. *Nature Reviews Genetics* 11:597–610. doi: 10.1038/nrg2843
 222. Bak M, Silahatoglu A, Moller M, et al. (2008) MicroRNA expression in the adult mouse central nervous system. *RNA* 14:432–444. doi: 10.1261/rna.783108
 223. Mercer TR, Dinger ME, Sunkin SM, et al. (2008) Specific expression of long noncoding RNAs in the mouse brain. *Proc Natl Acad Sci USA* 105:716–721. doi: 10.1073/pnas.0706729105
 224. Kapsimali M, Kloosterman WP, de Bruijn E, et al. (2007) MicroRNAs show a wide diversity of expression profiles in the developing and mature central nervous system. *Genome Biol* 8:R173. doi: 10.1186/gb-2007-8-8-r173
 225. Visvanathan J, Lee S, Lee B, et al. (2007) The microRNA miR-124 antagonizes the anti-neural REST/SCP1 pathway during embryonic CNS development. *Genes & Development* 21:744–749. doi: 10.1101/gad.1519107
 226. Rege SD, Geetha T, Pondugula SR, et al. (2013) Noncoding RNAs in Neurodegenerative Diseases. *International Scholarly Research Notices* 2013:1–5. doi: 10.1155/2013/375852
 227. Smith P, Hashimi AI A, Girard J, et al. (2011) In vivo regulation of amyloid precursor protein neuronal splicing by microRNAs. *Journal of Neurochemistry* 116:240–247. doi: 10.1111/j.1471-4159.2010.07097.x
 228. Lukiw WJ (2007) Micro-RNA speciation in fetal, adult and Alzheimer's disease hippocampus. *Neuroreport* 18:297–300. doi: 10.1097/WNR.0b013e3280148e8b
 229. Hébert SS, Horré K, Nicolai L, et al. (2008) Loss of microRNA cluster miR-29a/b-1 in sporadic Alzheimer's disease correlates with increased BACE1/beta-secretase expression. *Proc Natl Acad Sci USA* 105:6415–6420. doi: 10.1073/pnas.0710263105
 230. Wang X, Liu P, Zhu H, et al. (2009) miR-34a, a microRNA up-regulated in a double transgenic mouse model of Alzheimer's disease, inhibits bcl2 translation. *Brain Research Bulletin* 80:268–

273. doi: 10.1016/j.brainresbull.2009.08.006
231. Miñones-Moyano E, Porta S, Escaramís G, et al. (2011) MicroRNA profiling of Parkinson's disease brains identifies early downregulation of miR-34b/c which modulate mitochondrial function. *Human Molecular Genetics* 20:3067–3078. doi: 10.1093/hmg/ddr210
232. Kabaria S, Choi DC, Chaudhuri AD, et al. (2015) Inhibition of miR-34b and miR-34c enhances α -synuclein expression in Parkinson's disease. *FEBS Lett* 589:319–325. doi: 10.1016/j.febslet.2014.12.014
233. Lee S-T, Chu K, Im W-S, et al. (2011) Altered microRNA regulation in Huntington's disease models. *Experimental Neurology* 227:172–179. doi: 10.1016/j.expneurol.2010.10.012
234. Marti E, Pantano L, Banez-Coronel M, et al. (2010) A myriad of miRNA variants in control and Huntington's disease brain regions detected by massively parallel sequencing. *Nucleic Acids Research* 38:7219–7235. doi: 10.1093/nar/gkq575
235. Packer AN, Xing Y, Harper SQ, et al. (2008) The Bifunctional microRNA miR-9/miR-9* Regulates REST and CoREST and Is Downregulated in Huntington's Disease. *J Neurosci* 28:14341–14346. doi: 10.1523/J.Neurosci.2390-08.2008
236. Lee Y, Samaco RC, Gatchel JR, et al. (2008) miR-19, miR-101 and miR-130 co-regulate ATXN1 levels to potentially modulate SCA1 pathogenesis. *Nat Neurosci* 11:1137–1139. doi: 10.1038/nn.2183
237. Novoa EM, Ribas de Pouplana L (2012) Speeding with control: codon usage, tRNAs, and ribosomes. *Trends in Genetics* 28:574–581. doi: 10.1016/j.tig.2012.07.006
238. Goodenbour JM, Pan T (2006) Diversity of tRNA genes in eukaryotes. *Nucleic Acids Research* 34:6137–6146. doi: 10.1093/nar/gkl725
239. Ishimura R, Nagy G, Dotu I, et al. (2014) RNA function. Ribosome stalling induced by mutation of a CNS-specific tRNA causes neurodegeneration. *Science* 345:455–459. doi: 10.1126/science.1249749
240. Borthwick GM, Taylor RW, Walls TJ, et al. (2006) Motor neuron disease in a patient with a mitochondrial tRNA^{Ala} mutation. *Ann Neurol* 59:570–574. doi: 10.1002/ana.20758
241. Rudnik-Schöneborn S, Barth PG, Zerres K (2014) Pontocerebellar hypoplasia. *Am J Med Genet C Semin Med Genet* 166C:173–183. doi: 10.1002/ajmg.c.31403
242. Budde BS, Namavar Y, Barth PG, et al. (2008) tRNA splicing endonuclease mutations cause pontocerebellar hypoplasia. *Nat Genet* 40:1113–1118. doi: 10.1038/ng.204
243. Trotta CR, Paushkin SV, Patel M, et al. (2006) Cleavage of pre-tRNAs by the splicing endonuclease requires a composite active site. *Nature* 441:375–377. doi: 10.1038/nature04741
244. Kasher PR, Namavar Y, van Tijn P, et al. (2011) Impairment of the tRNA-splicing endonuclease subunit 54 (tsen54) gene causes neurological abnormalities and larval death in zebrafish models of pontocerebellar hypoplasia. *Human Molecular Genetics* 20:1574–1584. doi: 10.1093/hmg/ddr034
245. Edvardson S, Shaag A, Kolesnikova O, et al. (2007) Deleterious mutation in the mitochondrial arginyl-transfer RNA synthetase gene is associated with pontocerebellar hypoplasia. *Am J Hum Genet* 81:857–862. doi: 10.1086/521227
246. Hanada T, Weitzer S, Mair B, et al. (2013) CLP1 links tRNA metabolism to progressive motor-neuron loss. *Nature* 1–7. doi: 10.1038/nature11923
247. Karaca E, Weitzer S, Pehlivan D, et al. (2014) Human CLP1 Mutations Alter tRNA Biogenesis, Affecting Both Peripheral and Central Nervous System Function. *Cell* 157:636–650. doi: 10.1016/j.cell.2014.02.058
248. Schaffer AE, Eggens VRC, Caglayan AO, et al. (2014) CLP1 Founder Mutation Links tRNA Splicing and Maturation to Cerebellar Development and Neurodegeneration. *Cell* 157:651–663. doi: 10.1016/j.cell.2014.03.049
249. Qureshi IA, Mattick JS, Mehler MF (2010) Long non-coding RNAs in nervous system function and disease. *Brain Research* 1338:20–35. doi: 10.1016/j.brainres.2010.03.110
250. Faghihi MA, Zhang M, Huang J, et al. (2010) Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol* 11:R56. doi: 10.1186/gb-2010-11-5-r56
251. Tan JY, Vance KW, Varela MA, et al. (2014) Cross-talking noncoding RNAs contribute to cell-specific neurodegeneration in SCA7. *Nat Struct Mol Biol* 21:955–961. doi: 10.1038/nsmb.2902
252. Francelle L, Galvan L, Gaillard M-C, et al. (2015) The striatal long noncoding RNA Abhd11os is neuroprotective against an N-terminal fragment of mutant huntingtin in vivo. *Neurobiology of Aging* 36:1601.e7–16. doi: 10.1016/j.neurobiolaging.2014.11.014
253. Guthrie C, Patterson B (1988) Spliceosomal

- snRNAs. *Annu Rev Genet* 22:387–419. doi: 10.1146/annurev.ge.22.120188.002131
254. Maniatis T, Reed R (1987) The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. *Nature* 325:673–678. doi: 10.1038/325673a0
255. Zhang Z, Lotti F, Dittmar K, et al. (2008) SMN deficiency causes tissue-specific perturbations in the repertoire of snRNAs and widespread defects in splicing. *Cell* 133:585–600. doi: 10.1016/j.cell.2008.03.031
256. Jia Y, Mu JC, Ackerman SL (2012) Mutation of a U2 snRNA Gene Causes Global Disruption of Alternative Splicing and Neurodegeneration. *Cell* 148:296–308. doi: 10.1016/j.cell.2011.11.057





CHAPTER III

Genetic screens in *Caenorhabditis elegans* models for neurodegenerative diseases



Olga Sin^{1,2}, Helen Michels¹ & Ellen A.A. Nollen¹

¹ University of Groningen, University Medical Centre Groningen, European Research
Institute for the Biology of Aging, 9700 AD Groningen, The Netherlands

² Graduate Program in Areas of Basic and Applied Biology, Abel Salazar Biomedical
Sciences Institute, University of Porto, 4099-003 Porto, Portugal.

Abstract

Caenorhabditis elegans comprises unique features that make it an attractive model organism in diverse fields of biology. Genetic screens are powerful to identify genes and *C. elegans* can be customized to forward or reverse genetic screens and to establish gene function. These genetic screens can be applied to “humanized” models of *C. elegans* for neurodegenerative diseases, enabling for example the identification of genes involved in protein aggregation, one of the hallmarks of these diseases. In this review, we will describe the genetic screens employed in *C. elegans* and how these can be used to understand molecular processes involved in neurodegenerative and other human diseases.

1. Introduction

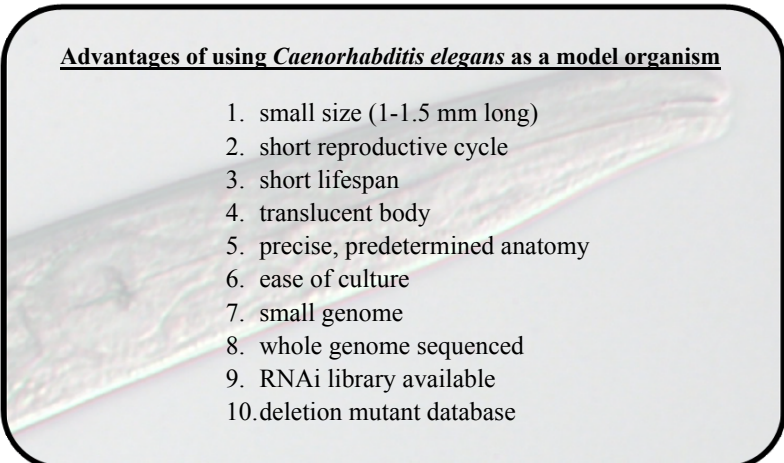
1.1 *Caenorhabditis elegans* as a model organism

Sydney Brenner first introduced the nematode *C. elegans* as a genetic model organism in 1965 and since then the model has been extensively used in very diverse fields of research, from developmental biology to ecotoxicology, aging and neuroscience [1]. This has resulted in several breakthroughs in biomedical science, which include the discovery of genetic regulators of programmed cell death, the use of the green fluorescent protein as a protein marker, and the discovery of RNA interference. Indeed, this nematode combines a number of characteristics that make it an advantageous model, anatomically and genetically, which are summarized in Box 1. Moreover, the characteristics of this invertebrate make it an easy experimental model to study biological processes in a relatively cheap, quick, and easy way.

C. elegans is a small, free-living nematode of about 1–1.5 mm in length that can be found in temperate soil environments feeding on different bacteria, including *Escherichia coli*. It exists in two sexual forms, as a hermaphrodite or as a male. The former is self-fertile, able to produce its own sperm and eggs and is the predominant adult form. Although males are rare (about 0.02%), their abundance in the offspring can be increased to 50% by mating with

Box 1

Advantages of using *Caenorhabditis elegans* as a model organism

- 
1. small size (1-1.5 mm long)
 2. short reproductive cycle
 3. short lifespan
 4. translucent body
 5. precise, predetermined anatomy
 6. ease of culture
 7. small genome
 8. whole genome sequenced
 9. RNAi library available
 10. deletion mutant database

hermaphrodites [2]. The length of the life cycle of wild type N2 *C. elegans* strains and its lifespan depends on the growth temperature. Grown at 20°C, hermaphrodites usually lay 300–350 eggs and once the eggs hatch, it takes about three days to develop from a larva to an adult. The average lifespan of this organism can vary between 18 and 20 days [3,4]. At higher temperatures, the life cycle is shortened and the lifespan decreased. One major advantage of *C. elegans* is that it has a well-dissected and predetermined anatomy. The adult hermaphrodite has exactly 959 somatic cells and 302 neurons [1,5,6]. Its transparent body enables one to easily follow cell fate or expression of fluorescently tagged proteins of interest in the living animal. Moreover, *C. elegans* was the first multicellular organism to have the complete genome sequenced and this gave rise to several databases and resources that are currently available online for the scientific community ([7], see “Online links” at the end of this article).

Genetic screens are widely used in *C. elegans* to discover gene function. It can be easily applied to discover which gene mutations are responsible for a specific phenotype of interest (forward genetics) or, conversely, the gene function can be purposely altered to assess what is the consequence in terms of development, behavior or alterations in specific biological processes (reverse genetics). The two major genetic screens employed are ethyl methane sulfate (EMS) screens and (genome-wide) RNAi screens. They have been fundamental not only to dissect nematode genetics but also to identify genes involved in aging, development, DNA damage response, and signal transduction, amongst other biological processes [8–13].

1.2 *C. elegans* homology to humans

For the scope of this review, we shall explore the rationale and the basic procedures for both methods not only highlighting their advantages but also pinpointing the drawbacks. Next, we will explore how genetic screens can help us gain insight into the molecular and cellular mechanisms of human diseases. Specifically, we will focus on the application of genetic screens to discover potential disease-modifier genes by exemplifying studies on *C. elegans* models for neurodegenerative diseases.

There are a significant number of proteins that are evolutionary conserved between *C. elegans* and humans. At the time that the *C. elegans* genome sequencing was complete, 36% of *C. elegans* proteins (from a set composed of 18,891 protein sequences) were found to have homologs in humans (set composed of 4979 protein sequences), by pairwise comparison (The *C. elegans* Sequencing Consortium, 1998 [7]). Thereafter, this percentage was increased to 83% due to the much larger human gene dataset available to perform the comparison [14]. A more recent study estimated that 38% of the 20,250 *C. elegans* protein-coding genes had unique corresponding functional orthologs in humans (7663 unique hits) [15]. In a nutshell, biological processes unraveled in the invertebrate *C. elegans* can provide insight into human biology.

2. Genetic screens

Genetic screens in *C. elegans* are well-established and commonly used to assess gene function in any biological process of interest. High-throughput (semi-) automatized setups and screening methods enable hundreds of parallel experiments in microtiter plates. In a screen, wild type animals are mutagenized or treated with RNAi and then scored for phenotypical changes. Below we describe two types of genetic screens that are most frequently used: EMS mutagenesis and RNA interference (RNAi). The characteristics of both type of screens are summarized in Table 1.

Table 1. Features of EMS mutagenesis versus RNA interference.

EMS mutagenesis	RNA interference
Inactivation or alteration of gene function	Reduction or depletion of gene function
Requires identification of gene mutation	Candidate gene is known
Permanent mutation	Possible to select developmental stage for depletion No effects on embryos in the first generation
Can select for non-essential genes	Can identify roles of essential genes in a post-developmental process Limited penetrance to neurons Limited efficiency if the protein that is encoded by the targeted gene is very stable

2.1 EMS mutagenesis

The most commonly used method to mutate the genome of *C. elegans* is the treatment with EMS. The mutagen induces mutations in the sperm and oocytes of hermaphrodites. Sydney Brenner tested systematically different mutagens, but researchers are mostly using EMS because of its relatively low toxicity and relatively good efficiency (summarized in [16]). The hermaphroditism of *C. elegans* allows easy maintenance of a mutation, as a homozygous worm will pass it to all the progeny through self-fertilization.

Mutations can be identified using a simple F2 screen firstly described by Brenner in 1974 [1]. Thousands of copies of any particular gene can be analyzed in a typical EMS screen. The frequency of a null mutation at any particular locus of the genome is one for every 2000 copies by using standard concentrations (50 mM) of the mutagen. That means that one can expect to identify 6 mutations per particular gene in a typical experiment of 12,000 haploid genomes (reviewed in [16]). The mutagenized worms are placed on Petri dishes and grown for two generations to produce homozygous mutants (Fig. 1). Worms from the F2 generation showing a specific phenotype of interest are further singled to new plates to determine whether the phenotype is transmitted to the next generation.

Once a worm with a specific phenotype is isolated, the responsible mutation needs to be identified. By using single nucleotide polymorphisms (SNPs) of the Hawaiian wild type strain in comparison to the Bristol strain (natural variation wild type) it is possible to map a mutation first to a certain chromosome [17] and then in several steps to a specific region on that chromosome. When a mutation is mapped to a gene region, sequencing or the specific knockdown of every single gene in that area by RNAi can be used to identify the mutated gene. The development of new sequencing methods like deep sequencing in the last decade facilitates the identification of mutations and can save laborious fine mapping [18–20]. It is important to keep in mind that an isolated, mutated animal can have several mutations at different loci. For further interpretation, it is therefore necessary to backcross the animals several times with wild type strains. The importance for controlling the genetic background was shown by

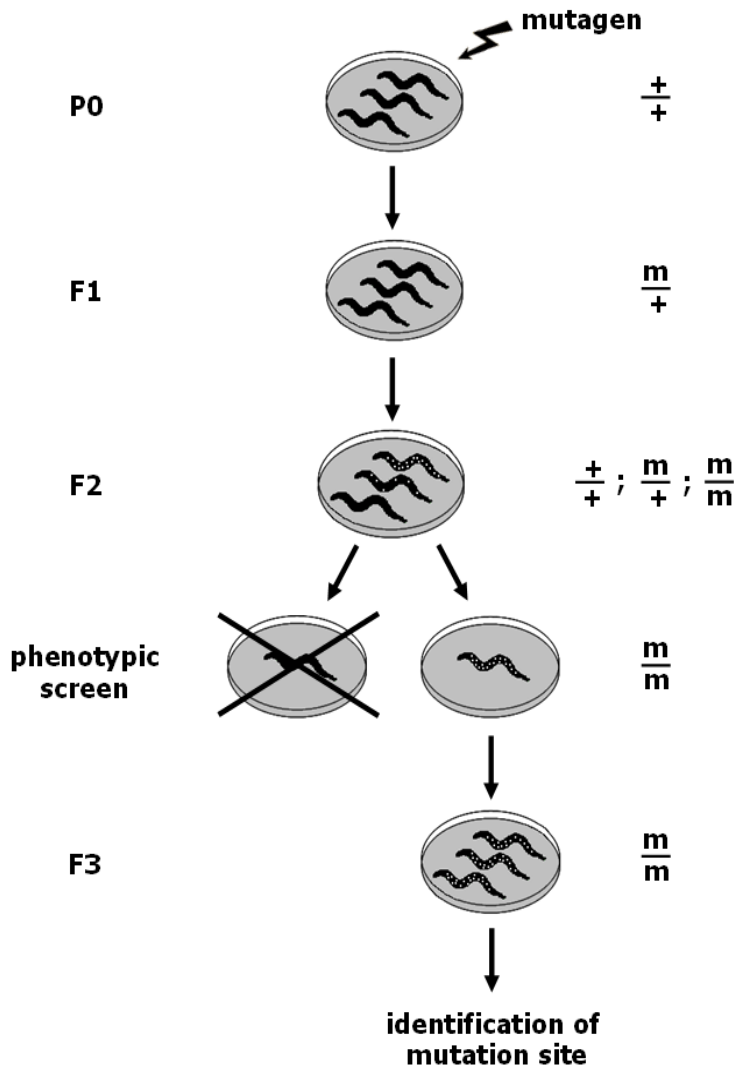


Figure 1. High-throughput mutagenesis screen in *C. elegans*: Animals (P0) are treated with a mutagen (e.g. EMS) to produce progeny (F1) containing mutations in alleles of different genes. These animals get progeny (F2) of different genotypes by self-fertilization which are further scored for a specific phenotype. Positive evaluated animals are singled to validate a homozygous mutation by breeding through of the phenotype to the next generation (F3). Once a mutant is isolated, the mutation site needs to be mapped to a specific genomic site to continue with functional studies (adapted from [16]).

Burnett and colleagues. They demonstrated that a described lifespan extension [21] by overexpression of SIR-2 disappeared after several backcrossings [22]. Tissenbaum and Guarente further showed that the overexpression of SIR-2.1 slightly increases lifespan but to a much lesser extent than the transgenic animals used in their first publication [23]. Deep sequencing methods can be

used to monitor mutations in the background. Preferentially, one should use independent mutant or deletion alleles of a gene to confirm results.

In the first EMS screens, 619 mutants were identified with visible phenotypes especially from the uncoordinated class [1]. This group of genes impairs wild type movements when mutated. Under laboratory conditions proper moving is not essential as food is plentiful and sex is dispensable so therefore maintenance and characterization of mutants that may not survive in non-laboratory conditions are possible. Many of these mutants have revealed important information about molecules and mechanisms involved in human disease. For example, one gene of the uncoordinated class is *unc-2* (uncoordinated 2) and encodes for a homolog of the voltage-sensitive calcium-channel alpha-1 subunit (human P/Q calcium channel CACNA1A) [24]. Missense mutations in the CACNA1A calcium channel in humans are associated with a rare form of migraine [25], which is often associated with low levels of serotonin [26]. *unc-2* mutants show neuronal migration defects similar to serotonin-deficient mutants [27] and UNC-2 is required for the desensitization to the two neurotransmitters dopamine and serotonin [28]. Studies using *C. elegans* showed that UNC-2 interacts with the transforming growth factor (TGF)- β , a pathway that is required for movements through regulation of serotonin levels probably through the modulation of the expression of *tph-1* (tryptophan hydroxylase), the enzyme that converts tryptophan into serotonin [29]. There are elevated levels of TGF- β 1 in migraine patients compared to those of pain-free individuals [30].

In addition to mutations in the uncoordinated class, Brenner also identified mutants with aberrant appearance like animals with small bodies, blistered cuticles, twitching muscles, rolling locomotion, long bodies, dumpy bodies, forked heads or bent heads [1].

EMS screens are often used to identify different mutations with the same phenotype to further investigate if those genes function in the same processes. Using this approach, John Sulston and H. Robert Horvitz searched e.g. for mutants that show defects in the differentiation of a vulva from epidermal cells [31]. Molecular follow-up studies revealed that animals that lack a vulva had

mutations in two signaling pathways: the epidermal growth factor (EGF)/RAS pathway and the Notch signaling pathway (reviewed in [32–34]), both having major roles in cell fate determination. These studies in *C. elegans* have increased the understanding of these molecular pathways involved in oncogenesis in humans (reviewed in [35,36]).

Another possibility to find genes of the same genetic pathways are enhancer or suppressor screens. In this case the mutagenesis occurs on a non-wild type strain whose genetic composition is known and causes a defined phenotype. Like this, one can screen for mutations in this genetic background that enhance or suppress (reverse) that phenotype. With this approach one is able to show that two genes not only act in the same pathway but also their hierarchy which means that one is acting upstream of the other (summarized in [16]). However, one should still keep in mind that it might also be possible that some proteins result in the same phenotype when mutated even though they do not necessarily function in the same pathway.

Although EMS mutagenesis is a powerful tool to generate a high number of mutations and high-throughput screens to identify mutants with a specific phenotype it also has some limitations. It has been estimated that about 30% of the genes in *C. elegans* can be mutated to a visible phenotype [37] (some mutations might result e.g. in a lethal phenotype as it is the case for a number of developmental genes) and it needs to be mentioned that the identification of the same mutations which indicates a saturation of the screen is no guarantee that some other genes might be missed in this screen. High-throughput screens are only a starting point for further detailed experiments at molecular levels.

2.2. RNA interference

RNA interference was first discovered and investigated in *C. elegans* and published in 1998 by Andrew Fire et al. [38] (Nobel Prize in Physiology and Medicine in 2006). The discovery of dsRNA-mediated gene silencing has revolutionized genetic studies in *C. elegans*, as well as in other model organisms. Similar to EMS screens, RNAi screens can be used to identify genes that, when depleted, result in a certain phenotype or enhance or suppress a mutant phenotype.

RNAi in *C. elegans* is systemic, which, to date, is not the case for any other animal models. Therefore it is sufficient to introduce dsRNA into one specific tissue to get RNA silencing also in distant cells because of an amplification process called transitive RNAi [39]. This systemic effect is advantageous for large-scale genome-wide RNAi screens in *C. elegans*.

There are different methods that describe how to silence gene expression in *C. elegans*. The dsRNA can be delivered into the worm by (1) injection into any region [38], (2) feeding with dsRNA-producing bacteria [40], (3) soaking in dsRNA [41] or (4) in vivo production of dsRNA from transgenes under the control of specific promoters [42]. Cell-specific factors seem to regulate thereby the entry and export of dsRNA [39,43,44]. Some cell types (e.g. neurons) seem not to respond well to systemically delivered RNAi [42]. The use of RNAi enhanced mutants (e.g. *eri-1* mutant or mutants of the retinoblastoma pathway that are described to enhance RNAi especially in nervous tissue) might circumvent this problem [45,46]. In addition, Calixto and colleagues generated transgenic animals overexpressing the transmembrane protein SID-1 which is an essential component for systemic RNAi in the neurons. This modification increased the response to dsRNA delivered by feeding. It seemed that the expression of SID-1 in the neurons, on the other hand, decreased the RNAi effect on non-neuronal cells which might be useful for studying the function of essential genes in the neurons. This effect could be even increased when using a *sid-1* mutant background [47]. Durieux and colleagues further used this mutant that is insensitive for systemic RNAi to investigate the knockdown of one of the cytochrome c oxidase-1 subunits in specific tissues by controlling the expression of dsRNA via tissue-specific promoters [48]. The tissue-specific expression of SID-1 (not only in neurons) in a *sid-1* knockout background probably also enables to study tissue-specific effects especially of essential genes.

Especially the possibility to feed animals with dsRNA-producing bacteria enables to perform high-throughput RNAi screens in *C. elegans* (Fig. 2) [49,50]. For efficient induction of RNA interference the choice of the dsRNA-coding region is essential. In *C. elegans*, long dsRNA fragments (more than 100 bp) trigger gene silencing via RNAi. For most genes dsRNA is about 200–1000 nucleotides or even longer and covers exon-regions of the targeted gene. The

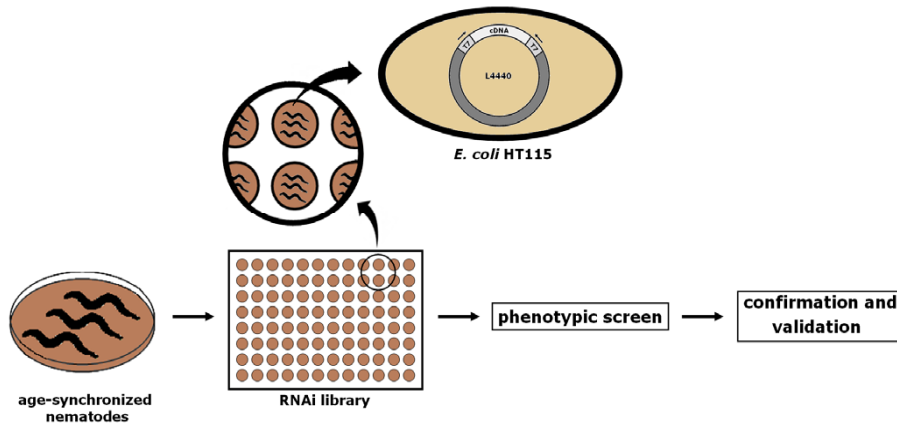


Figure 2. High-throughput RNAi screen in *C. elegans*: Age-synchronized animals are transferred to microtiter plates containing different clones of HT115 *E. coli* bacteria. Every clone produces a specific dsRNA which is taken up by the nematodes and induces a knockdown of the corresponding gene. Positive hits in the phenotypic screen are finally confirmed by sequencing the bacterial clone and repeating the specific knockdown in single experiments. Starting point of the RNAi feeding (possible at any developmental stage) and time point of the phenotypic scoring depend on the experiment setup

fragment should only target one gene. Once the coding region is chosen it can be cloned into a specific vector encoding the production of the specific dsRNA (summarized in [51]). The L4440 vector contains two bacteriophage T7 RNA polymerase promoters flanking the multiple cloning site in which the cDNA of a specific gene has been inserted. The construct can be transformed into *E. coli* strain HT115 (ED3). This strain is deficient for the bacterial RNA polymerase III and its production of bacteriophage T7 polymerase from the construct can be induced by the addition of isopropyl β -D-1-thiogalactopyranoside (IPTG). The bacteria are then synthesizing two complementary RNA strands that form a duplex RNA which can mediate RNAi [38].

RNAi libraries are commercially available which includes one library of bacteria clones containing cDNAs of 17,575 genes which represents about 87% of the *C. elegans* genome [49] and one library including clones of 11,800 *C. elegans* genes [52]. Positive scored clones subsequently can be sequenced to confirm that they target the predicted gene. To prevent any further off-target effects of the dsRNA and therefore false-positive results, one should consider generating a second RNAi construct targeting the mRNA of the same gene [52]. Besides the coding regions, the 3'UTR of mRNA might as well be a suitable target as RNA localization elements for the transport of the mRNA or regulation elements of

eukaryotic gene expression are typically located in this region (summarized in [53,54]).

Dissolving adult hermaphrodites with hypochlorite (bleach) will yield only fertilized eggs and can be used to age-synchronize the animals for a screen. Gene knockdown by RNAi can be induced at different developmental stages in contrast to EMS mutagenesis, which generates stable mutations that are present at all stages. Thus, the examination of the function of a gene that is transcribed at different developmental stages is possible. This is especially interesting when an active gene is essential at early developmental stages [55]. To investigate the effect of gene depletion by RNAi during embryonic development it is necessary to feed the parental worm with the specific bacterial strain.

Another possibility to study the effect of gene depletion at a certain timepoint was previously described by Calixto and colleagues, by performing a temperature-sensitive conditional knockdown. They could induce the knockdown of a gene by controlling via temperature the expression of RDE-1, a *C. elegans* argonaut protein which is required for RNA interference. This resulted in active RNAi at 15°C but not at 25°C. Furthermore, they observed that the switching ON and OFF is much faster than transferring animals from RNAi-mediating bacteria to non-RNAi inducing bacteria and vice-versa [56].

The dilution of bacteria containing a specific RNAi construct with non-RNAi mediating bacteria may decrease the efficiency of knocking down a certain gene. In that case of mild RNA interference, lethality effects and other very strong phenotypes are reduced and it might still be possible to study the function of these special genes.

The target of RNAi is known. This is one major difference to EMS mutagenesis that cannot be directed to specific genes. Therefore, besides genome-wide RNAi screens, one can also screen in a smaller subset of candidate genes for example based on previous microarray data, GWAS data, interactome studies, etc. Colaiácovo and colleagues performed for example a RNAi screen to check for germline phenotypes in a subset of genes that were generated by

a previous microarray analysis of Reinke et al. that was focusing on germline-enriched gene expression [57,58]. In another screen we were looking for genes that, when knocked down, increased the number of alpha-synuclein inclusions in a Parkinson's disease *C. elegans* model [59]. This group of 80 genes was then further used for a second RNAi screen in order to find candidates that, when knocked down, induced motility changes in the disease background [60].

RNAi efficiency of bacterial clones in the library can differ. Whereas some dsRNAs induce gene silencing closely to a knockout of a gene, others only generate a mild knockdown. One should always be aware that RNAi is only silencing gene activity and that it is not a full knockout of a gene. It is estimated that about 10–30% of candidates are scored as false negatives as the RNAi is not efficient enough to result in an obvious phenotype. On the other hand the percentage of false positives is relatively low (0.4%) [61]. It is also important to keep in mind that RNA interference is acting at the mRNA level and therefore only influencing the expression of a protein. That means that the stability of a protein is highly influencing the RNAi effect as already generated proteins and their activity are not affected anymore.

Results can also differ from one experiment to the other using the same bacteria clone to silence a specific gene. For example the freshness of the material can be crucial (IPTG, Ampicillin, RNAi construct containing bacteria) [51]. In contrast, a knockout mutant e.g. by EMS mutagenesis has the advantage that it results in a stable genotype. Results of the RNAi screen should therefore be confirmed by single experiments with the candidate genes, preferentially with a gene mutant strain.

A clear phenotype for scoring is mandatory for any successful screen. An obvious easy-to-recognize phenotype as well as automation of scoring facilitates the screening process.

The small size of the animals, the variety of simple phenotypes that are often results of one single gene disruption or silencing, the hermaphroditic reproduction, the homology to higher organisms (see above) and the knowledge of the *C. elegans* genome, cell-distribution and nematode anatomy

make this animal an optimal model organism to identify the function of genes via any kind of high-throughput screen.

3. From genome to function: what have genetic screens taught us?

One of the advantages of *C. elegans* is that it is amenable to generate “humanized” models of human diseases. For the purpose of this review, we will describe as an example *C. elegans* models of neurodegenerative diseases. Neuropathological hallmarks found in the human brain can be successfully recapitulated in the nematode, such as protein aggregation [62]. Indeed, one of the common features in neurodegenerative diseases is the presence of protein aggregates in the brains of affected patients. These structures originate from protein misfolding and aggregation of so-called “aggregation-prone proteins”. To name a few, these can be the amyloid-beta in Alzheimer’s disease (AD), mutant huntingtin in Huntington’s disease (HD) and alpha-synuclein in Parkinson’s disease (PD) [63]. By mechanisms that are still to be unraveled, these aggregation-prone proteins adopt a distinct conformation, which is thought to be a toxic gain-of-function [64,65]. The general understanding is that aggregation (or inclusion formation) renders cellular protection by sequestering misfolded proteins, therefore preventing potentially toxic protein–protein interactions [65,66].

Several nematode models have been generated to recapitulate molecular aspects of diseases, including HD, PD, AD and muscular dystrophy [59,67–72]. Although they do not feature clinical aspects of the disease, they provide the means to understand the molecular mechanisms in these diseases. Genetic screens performed in some of these models represent quick, unbiased methods

Table 2. Summary of genetic screens performed in *C. elegans* models of neurodegenerative diseases.

Model	Transgene	Tissue	Genetic screen	Modifier hits	Mammalian ortholog	Cellular process(es)	Transposed to	Reference
PolyQ diseases	Q35-YFP	Body wall muscle	Genome-wide RNAi screen	186 genes	n.a.	RNA synthesis and processing, protein synthesis, folding, transport, degradation, components of the proteasome	n.a.	[83]
PolyQ diseases	Q35-YFP	Body wall muscle	Genome-wide RNAi screen	88 genes	n.a.	Cell cycle, cell structure, protein transport and energy and metabolism	n.a.	[84]
PolyQ diseases	Q40-YFP	Body wall muscle	EMS screen	MOAG-4	SERF1A and SERF2	Unknown	Human cell models	[76]
PD	alpha-synuclein-YFP	Body wall muscle	Genome-wide RNAi screen	80 genes	n.a.	Vesicle trafficking, lipid metabolism, lifespan	n.a.	[59]
PD	alpha-synuclein-YFP	Body wall muscle	RNAi	TDO-2	TDO	Tryptophan metabolism	n.a.	[60]
PD	alpha-synuclein-GFP	Body wall muscle and DA neurons	Hypothesis-based RNAi screen	ATGR7	ATG7	Autophagy	Mice	[69, 78]
PD	alpha-synuclein (WT, A53T, A30P)	Whole nervous system	RNAi screen	APA-2; APS-2	n.a.	Vesicular trafficking	n.a.	[70]
AD	tau	Whole nervous system	Genome-wide RNAi screen	60 genes	38 have human homologs	Kinases and phosphatases; protein folding, stress response and degradation; transcription; proteolysis; neurotransmission and signaling; neuronal regeneration	n.a.	[75]

that have enabled insights into the underlying mechanisms of neurodegeneration. Indeed, many of the disease modifiers discovered in *C. elegans* were found to be reproducible in human cell-based models and other animal models such as mice, strengthening the validity of using this small organism to study complex human diseases, as summarized in Table 2 [69,73–78].

3.1. *C. elegans* models for polyglutamine diseases

Polyglutamine diseases comprise a subset of neurodegenerative disorders that include HD, spinocerebellar ataxias (– 1, – 2, – 6, – 7, –17), Machado–Joseph disease (also known as spinocerebellar ataxia 3) and spinobulbar muscular atrophy [79]. The common characteristic of polyglutamine diseases is an abnormal expansion of CAG triplets (which encode glutamine) in the coding region of the disease gene. Although the length of the CAG repeat may vary from individual to individual, the threshold to develop disease is around 40 CAG repeats (except for SCA6), which cause a polyglutamine expansion in the protein that is prone to aggregate. The larger the CAG repeat the earlier onset will occur and the more severe the disease phenotype will be. A more detailed and complete information on polyglutamine diseases is reviewed elsewhere [79]. In *C. elegans*, several different models have successfully recapitulated protein aggregation. Similarly to what occurs in humans, the length of the CAG repeats also determines the aggregation phenotype in *C. elegans*. At least three models have been generated to induce polyglutamine-associated toxicity in neurons by expressing expanded polyglutamine stretches in ASH sensory neurons, touch receptor neurons or the entire nervous system of *C. elegans* [68,80–82]. Polyglutamine aggregation has been modeled in the body wall muscle cells of *C. elegans* [68]. In this model, expanded polyglutamine stretches are fused to a yellow fluorescent protein (YFP) under the *unc-54* promoter, which is specific to the body wall muscle. The aggregation and toxicity phenotype is polyglutamine length-dependent. As the animal ages, the accumulation of protein aggregates increases, which is associated with toxicity [68]. This model has been widely used for genetic screens to discover enhancers and/or suppressors of polyglutamine proteotoxicity. Two genome-wide RNAi screens revealed modifier genes and classified them according to their biological function [83,84]. In the first screen, Q35 animals were fed with

dsRNA-producing bacteria and scored for genes that, when downregulated, provoked premature polyglutamine aggregation [83]. The major functional classes included RNA synthesis and processing, protein synthesis, folding, transport and degradation and components of the proteasome. In the second screen, the authors sought for genes that drive aggregation in Q35 animal and therefore the selection was made for genes that suppressed polyglutamine-induced aggregation when downregulated [84]. With this study, a new subset of modifier genes was recently found to belong to broader biological functions, namely cell cycle, cell structure, protein transport and energy and metabolism [83,84]. Therefore, proteotoxicity is not derived only from protein-related processes but rather a more diverse spectrum of biological functions that also have an effect on protein misfolding and aggregation. Interestingly, nine of these recently identified modifier genes were able to fold misfolded proteins back into the native state when constitutively expressed in misfolding mutants [84].

Forward genetics have also been used to identify modifiers of proteotoxicity. One such screen consisted in treating Q40-expressing worms with EMS. The aim was to find positive regulators of aggregation by selecting genes that suppressed protein aggregation when chemically mutated by EMS. The screen revealed MOAG-4 (modifier of aggregation) as a general aggregation-promoting factor in polyglutamine, Parkinson's and Alzheimer's disease models [76]. Inactivating MOAG-4 alleviated from polyglutamine-induced aggregation and toxicity; moreover, this effect was functionally conserved in the human orthologs SERF1A and SERF2. A recent follow-up on one of these orthologs showed that SERF1A is a specific aggregation-promoting factor, since it was able to bind specifically to amyloidogenic proteins, including alpha-synuclein, prion protein, amyloid-beta and huntingtin, but not to non-amyloidogenic proteins [85].

Genetic screens have also been used to find regulators of proteotoxicity using the *C. elegans* neuronal system. An RNAi screen performed in a *C. elegans* model expressing 128 polyQ stretches in the touch receptor neurons resulted in 662 genes that either enhanced or suppressed neuron toxicity, as measured by loss of touch response [77]. Comparison of these disease modifier genes to gene

expression data in two mouse models of HD showed that there was an overlap of 49 genes that were dysregulated in the striatum of either model, emphasizing the power of using *C. elegans* to find novel regulators of proteotoxicity relevant in human diseases.

3.2. *C. elegans* models for Parkinson's disease

Parkinson's disease (PD) is the second most common neurodegenerative disease (after Alzheimer's disease) that affects 1% of the population over the age of 50. Clinically, it is characterized by resting tremors, rigidity, bradykinesia and postural instability [86,87]. The defects in the motor system result from the progressive loss of dopaminergic neurons in the *substantia nigra pars compacta* (SNpc), which project and innervate the neurons in the caudate and putamen. Consequently, there is a reduction of dopamine levels, which is the neurotransmitter that plays a role in the coordination of body movements. Besides motor disabilities, PD patients can experience non-motor symptoms such as autonomic dysfunction, sleep disturbances and neuropsychiatric symptoms [88]. Most cases of PD are sporadic (about 95%) with unknown etiology. It has been suggested that disease can result from the accumulation of toxins (pesticides and heavy metals) over the years. Only 5% of PD has a familial origin and is associated with genetic mutations [88]. However, there are neuropathological hallmarks common to both sporadic and familial forms of PD. These are the loss of dopaminergic neurons in the SNpc, that result from the degeneration of the nigrostriatal pathway which leads to the motor symptoms described earlier as well as the formation of intraneuronal protein aggregates known as Lewy bodies and Lewy neurites in the surviving neurons, which contain alpha-synuclein.

Alpha-synuclein is a small (140 amino acids) natively soluble, monomeric protein that is predominantly expressed in the brain and is enriched in presynaptic terminals [89]. Although the precise function of this protein remains unclear, it is thought to be involved in the regulation of dopamine neurotransmission, vesicular trafficking and modulation of synaptic function and plasticity [90–92]. Three different mutations in the alpha-synuclein gene (A53T, A30P and E46K) cause autosomal-dominant PD [93–95] and genomic duplications and triplications of the gene have also been identified; suggesting

that overproduction of wild type alpha-synuclein is sufficient to cause disease [96,97].

Genetic screens performed with this model have been supporting an important relationship between alpha-synuclein and vesicle transport. The “humanized” model of *C. elegans* for PD expresses the human alpha-synuclein fused to YFP in the body wall muscle. Phenotypically, immobile YFP-positive foci can be seen in the muscle cells and these foci increase in number and correlate with age-dependent toxicity. An unbiased genome-wide RNAi screen with this model showed 80 modifier genes that, when suppressed, provoked premature alpha-synuclein inclusion formation [59]. A follow-up on those modifier genes revealed *tdo-2*, a gene involved in tryptophan degradation, as a general regulator of protein homeostasis during aging [60]. Moreover, 49 of the original 80 modifier genes had human homologs, which were enriched for genes related to vesicular trafficking functions. In another screen using a similar model, nematode genes orthologous to human familial PD genes were preselected to perform a hypothesis-based RNAi screen [69]. A subset of candidate genes from the initial screen was then further analyzed in another *C. elegans* model, expressing alpha-synuclein in the dopamine neurons, in order to assess their relevance at the neuronal level. This study revealed five candidate genes that were able to protect from alpha-synuclein-induced dopaminergic neurodegeneration. Again, the most representative class of genes here was associated with vesicular trafficking, with the exception of the autophagy-related gene *Atgr7*, of which the mammalian ortholog (*Atg7*) was previously implicated in neurodegeneration in mice [78]. Also, a serine/threonine kinase involved in axonal elongation, UNC-51, was found to be homologous to the previously associated risk factor ULK-2, as revealed by a genome-wide association study performed in PD patients [98]. Parallel to these findings, Kuwahara et al. were able to pinpoint two genes, *apa-2* and *aps-2*, that when knockdown by RNAi increase alpha-synuclein-induced neurotoxicity in a *C. elegans* model expressing the transgene in the whole nervous system [70]. These two genes encode for subunits of the AP-2 adaptor complex, which mediates the internalization of cargo into the cell from the extracellular space via clathrin-mediated endocytosis [99].

3.3. C. elegans models for Alzheimer's disease

Alzheimer's disease (AD) is the most prevalent neurodegenerative disease, which is predicted to affect 66 million people worldwide by 2030 [100]. It represents the most common form of dementia, leading to clinical symptoms such as memory loss and mood swings. Aging and lifestyle are risk factors for development of AD, but 70% of the cases are attributable to genetics [101]. The main neuropathological features are the presence of extracellular amyloid-beta plaques, which consist of an accumulation of aggregated amyloid-beta, and intraneuronal tangles of hyperphosphorylated tau. Mutations in several genes can lead to the development of AD, including mutations in genes encoding for the amyloid-precursor protein (APP), presenilin 1 (PSEN1) and presenilin 2 (PSEN2). These genes are part of the APP cleavage pathway and mutations in these genes promote the processing of APP towards the amyloidogenic pathway, promoting the formation of amyloid-beta. Amyloid-beta peptides can have different lengths, including 40 or 42 amino acids. Amyloid-beta 42 is the most common species found in the amyloid plaques, indicating its propensity to rapidly aggregate in comparison to amyloid-beta 40.

Tau is encoded by the microtubule-associated tau protein (MAPT) gene and predominantly expressed in the nervous system. As to its function, it is known to associate and stabilize microtubules. It has been already classified as one of the risk genes for developing AD by at least two independent studies [102,103].

There are several models in *C. elegans* that express either human amyloid-beta or tau. In the first case, the worms express amyloid-beta 3–42 in the body wall muscles which causes the progressive accumulation of amyloid-beta 3–42 in the muscle cells and paralysis, which worsens with aging [71,104]. There have been several variations to this model, either combined with inducible systems, driving expression in neurons or, more recently, expressing full-length amyloid-beta 1–42 [105–107]. On the other hand, tau-expressing models have been specific to neuronal cells and the phenotype is either worsening of uncoordinated movement or insensitivity to the touch response due to transgene expression [75,108,109].

Genetic screens in *C. elegans* models for AD have been scarce. So far, there has been no genetic screen performed in any of the models expressing amyloid-beta. There is only one report on genome-wide RNAi done in a tau-expressing model [75]. Sixty modifier genes were discovered to belong to several functional classes including, kinases, chaperones, proteases and phosphatases. Of these, 38 had homologs in humans but, more importantly, 6 had already been associated with disease, either in humans or other animal models. One of these modifiers was the nicotinic acetylcholine receptor alpha-7 (nAChR), a ligand-gated ion channel expressed in the human brain and known to contribute to tau phosphorylation [110].

4. Final considerations

Genetic screens are powerful means to find genes involved in a certain biological process of interest and their function. The fact that *C. elegans* is a tractable system to model human diseases further allows one to perform genome-wide screenings in a relatively quick and unbiased manner. Genetic screens can have two outcomes, both being equally informative. On the one hand, new genes are discovered and therefore novel pathways are implicated, giving fresh perspectives on the biological process being studied. On the other hand, genetic screens that reveal genes already known to be associated with disease strengthen the importance of those genes in pathogenesis. Many screens that start with a genome-wide approach end up with an extensive list of candidate genes that are classified according to their functional class. From here, a selection of these genes should be refined and prioritized in order to study further their individual contribution to pathogenesis. One of the critical points is considering those that might have significance at the mammalian level. Additionally, if the human gene can replace the function of the endogenous one, it demonstrates evolutionary conservation of function and enables one to extrapolate findings from small organisms to complex human diseases. It is, therefore, essential to validate the genes from the screen in higher organisms. However, it is also often that a screen might reveal genes that do not have a direct sequence homolog in mammals. Nevertheless, they may be indicators of other genes that may be functional orthologs or otherwise regulators of genes with a role in human disease (e.g. transcription factors).

Another contribution of genetic screens may be to provide novel targets for drug development [111]. Although not all features of a complex human disease are fully recapitulated in the nematode, one can argue that its simplicity can be advantageous. Especially, because analysis of the expression of the causative gene and its interactors or modifiers can be done without other confounding factors inherent to the complexity of the human biology.

Genome-wide association studies (GWAS) have gained importance in the last ten years, becoming one of the forefront strategies to find common genetic factors associated with susceptibility to develop disease. One challenge now is to establish the functional consequences of these genetic variations. Coupling genetic screens or candidate gene approach in *C. elegans* to find causative genes may represent a quick and inexpensive way to assess functional relevance of associated variations and consequently obtain concrete targets to act upon. For instance, a genome-wide toxicity screen in yeast revealed 6 modifiers of amyloid-beta toxicity that were previously identified as risk factors in GWAS [112]. Importantly, those modifiers were functionally conserved from yeast, to *C. elegans* and to rat. Another study by Shulman et al. showed, for the first time, a link between an AD risk factor and a causative gene by functional screening in the fly [113].

Although this review focused in *C. elegans* models for neurodegenerative diseases, it should be noted that *C. elegans* is a model organism for other human diseases as well. *C. elegans* has been used to model certain aspects of cancer, diabetes, obesity, polycystic kidney disease, muscular dystrophy and innate immunity, to name a few. A more complete view of these different disease models is summarized elsewhere [111].

All in all, genetic screens in small organisms such as *C. elegans* can not only aid to dissect fundamental biological questions but also have the versatility of being adapted to model complex human diseases, such as neurodegenerative diseases. Moreover, its attributes make it a tractable system to drug target discovery and compound screening, emphasizing the potential of this organism to extrapolate findings from small organisms to higher vertebrates.

Online links

- Textpresso, a full text literature searches of *C. elegans*
(<http://www.textpresso.org/>)
- Worm Interactome Database
(http://interactome.dfci.harvard.edu/C_elegans/index.php)
- The Caenorhabditis Genetic Center, with an extensive list of strains
(<http://www.cbs.umn.edu/CGC/>)
- Wormbase, a complete database of genetics, genomics and biology of *C. elegans* (www.wormbase.org)
- Wormbook, a comprehensive, open-access collection of original, peer-reviewed chapters covering topics related to the biology of *C. elegans* and other nematodes (<http://wormbook.org>)
- *C. elegans* Gene Knockout Consortium, which creates knockout strains

(<http://celeganskoconsortium.omrf.org/>)
- National Bioresource Project, which generates, collects, stores and distributes deletion mutants of *C. elegans*
(<http://www.shigen.nig.ac.jp/c.elegans/index.jsp>)
- Wormatlas, a database of behavioral and structural anatomy of *C. elegans*
(<http://www.wormatlas.org>)

Acknowledgements

This project was funded by the European Research Council (ERC) starting grant (to E.A.A.N.) and by Fundação para a Ciência e Tecnologia (SFRH/BPD/51009/2010) (to O.S.).

References

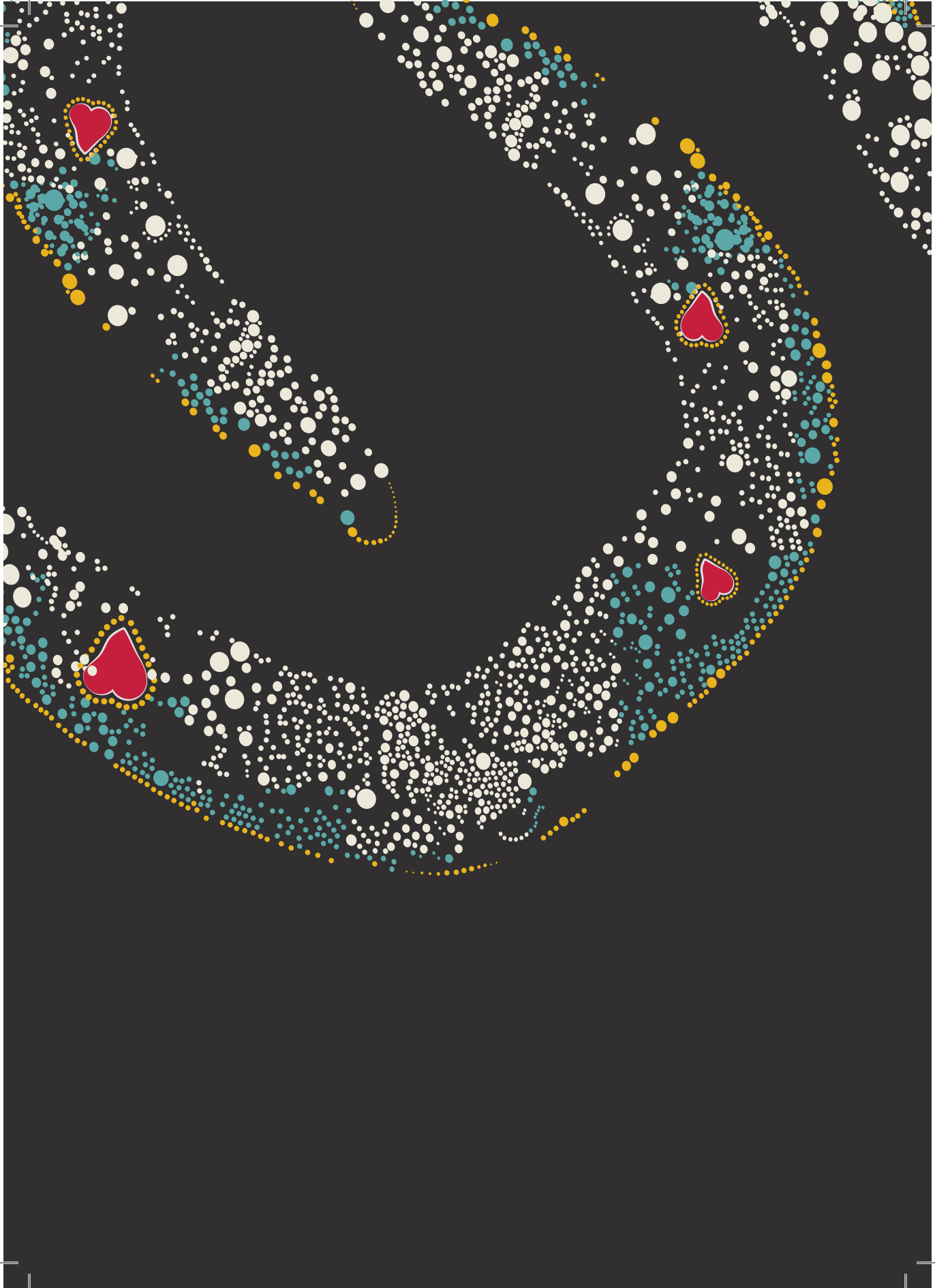
1. S. Brenner, The genetics of *Caenorhabditis elegans*, *Genetics* 77 (1974) 71–94.
2. D. Fay, Genetic mapping and manipulation: Chapter 1-Introduction and basics, in: *WormBook* (Ed.), The *C. elegans* Research Community, *WormBook*, February 17 2006, <http://dx.doi.org/10.1895/wormbook.1.90.1>, (<http://www.wormbook.org>).
3. S. Wolff, H. Ma, D. Burch, G.A. Maciel, T. Hunter, A. Dillin, SMK-1, an essential regulator of DAF-16-mediated longevity, *Cell* 124 (2006) 1039–1053.
4. T. Gidalevitz, N. Wang, T. Deravaj, J. Alexander-Floyd, R.I. Morimoto, Natural genetic variation determines susceptibility to aggregation or toxicity in a *C. elegans* model for polyglutamine disease, *BMC Biol.* 11 (2013) 100.
5. D.L. Riddle, T. Blumenthal, B.J. Meyer, et al., (Eds.), *C. elegans* II, 2nd edition, Section I, The Biological Model Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), 1997.
6. S.F. Gilbert, *Developmental Biology, Early Development of the Nematode Caenorhabditis elegans*, 6th edition, Sinauer Associates, Sunderland (MA), 2000.
7. The *C. elegans* sequencing consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology, *Science* 282 (1998) 2012–2018.
8. A. Dillin, A.L. Hsu, N. Arantes-Oliveira, J. Lehrer-Graiwer, H. Hsin, A.G. Fraser, R.S. Kamath, J. Ahringer, C. Kenyon, Rates of behavior and aging specified by mitochondrial function during development, *Science* 298 (2002) 2398–2401.
9. S.S. Lee, S. Kennedy, A.C. Tolonen, G. Ruvkun, DAF-16 target genes that control *C. elegans* lifespan and metabolism, *Science* 300 (2003) 644–647.
10. Y. Kim, H. Sun, Functional genomic approach to identify novel genes involved in the regulation of oxidative stress resistance and animal lifespan, *Aging Cell* 6 (2007) 489–503.
11. B. Hamilton, Y. Dong, M. Shindo, W. Liu, I. Odell, G. Ruvkun, S.S. Lee, A systematic RNAi screen for longevity genes in *C. elegans*, *Genes Dev.* 19 (2005) 1544–1555.
12. P. Zipperlen, A.G. Fraser, R.S. Kamath, M. Martinez-Campos, J. Ahringer, Roles for 147 embryonic lethal genes on *C. elegans* chromosome I identified by RNA interference and video microscopy, *EMBO J.* 20 (2001) 3984–3992.
13. C.D. Keating, N. Kriek, M. Daniels, N.R. Ashcroft, N.A. Hopper, E.J. Siney, L. Holden-Dye, J.F. Burke, Whole-genome analysis of 60 G protein-coupled receptors in *Caenorhabditis elegans* by gene knockout with RNAi, *Curr. Biol.* 13 (2003) 1715–1720.
14. C.H. Lai, C.Y. Chou, L.Y. Chang, C.S. Liu, W. Lin, Identification of novel human genes evolutionarily conserved in *Caenorhabditis elegans* by comparative proteomics, *Genome Res.* 10 (2000) 703–713.
15. D.D. Shaye, I. Greenwald, OrthoList: a compendium of *C. elegans* genes with human orthologs, *PLoS ONE* 6 (5) (2011) e20085.
16. Jorgensen, Mango, The art and design of genetic screens: *Caenorhabditis elegans*, *Nat. Rev. Genet.* 3 (5) (2002) 356–369.
17. Davis, et al., Rapid single nucleotide polymorphism mapping in *C. elegans*, *BMC Genomics* 6 (2005) 118.
18. Sarin, et al., *Caenorhabditis elegans* mutant allele identification by whole-genome sequencing, *Nat. Methods* 5 (10) (2008) 865.
19. Shen, et al., Comparing platforms for *C. elegans* mutant identification using high-throughput whole-genome sequencing, *PLoS ONE* 3 (12) (2008) e4012.
20. Bigelow, et al., A software, MAQGene, facilitating *C. elegans* whole genome sequence analysis for mutant identification, *Nat. Methods* 6 (8) (2009) 549.
21. Tissenbaum, Guarente, Increased dosage of sir-2 gene extends lifespan in *Caenorhabditis elegans*, *Nature* 410 (6825) (2001) 227–230.
22. Burnett, et al., Absence of effects of Sir-2 overexpression on lifespan in *C. elegans* and *Drosophila*, *Nature* 477 (7365) (2011) 482–485.
23. Viswanathan, Guarente, Regulation of *Caenorhabditis elegans* lifespan by sir-2.1 transgenes, *Nature* 477 (7365) (2011) E1–E2.
24. Bargmann, *Neurobiology of the Caenorhabditis elegans genome*, *Science* 282 (1998) 2028–2033.
25. Ophoff, et al., Familial hemiplegic migraine and episodic ataxia type-2 are caused by mutations in the Ca²⁺ channel gene CACNL1A4, *Cell* 87 (1996) 543–552.
26. Ferrari, et al., Serotonin metabolism in migraine,

- Neurology 39 (1989) 1239–1242.
27. Kindt, et al., Serotonin promotes G(o)-dependent neuronal migration in *Caenorhabditis elegans*, *Curr. Biol.* 12 (20) (2002) 1738–1747.
28. Schafer, Kenyon, A calcium-channel homologue required for adaptation to dopamine and serotonin in *Caenorhabditis elegans*, *Nature* 375 (6526) (1995) 73–78.
29. Estevez, et al., The voltage-gated calcium channel UNC-2 is involved in stress-mediated regulation of tryptophan hydroxylase, *J. Neurochem.* 88(2004)102–113.
30. Bo, et al., Cerebrospinal fluid cytokine levels in migraine, tension-type headache and cervicogenic headache, *Cephalalgia* 29 (3) (2009) 365–372.
31. Sulston, Horvitz, Abnormal cell lineages in mutants of the nematode *Caenorhabditis elegans*, *Dev. Biol.* 82 (1) (1981) 41–55.
32. Wang, Sternberg, et al., Pattern formation during *C. elegans* vulval induction, *Curr. Top. Dev. Biol.* 51 (2001) 189–220.
33. Beitel, et al., The *Caenorhabditis elegans* gene lin-1 encodes an ETS-domain protein and defines a branch of the vulval induction pathway, *Genes Dev.* 9 (1995) 3149–3162.
34. Jacobs, et al., Gain-of-function mutations in the *Caenorhabditis elegans* lin-1 ETS gene identify a C-terminal regulatory domain phosphorylated by ERK MAP Kinase, *Genetics* 149 (1998) 1809–1822.
35. Egan, Weinberg, The pathway to signal achievement, *Nature* 365 (6449) (1993) 781–783.
36. Pawson, Hunter, Signal transduction and growth control in normal and cancer cells, *Curr. Opin. Genet. Dev.* 4 (1) (1994) 1–4.
37. Waterston, et al., The Genome, in: Riddle, Blumenthal, Meyer, Priess (Eds.), *C. elegans* II, vol. 33, Cold Spring Harbor Laboratory Press, New York, 1997, pp. 23–45.
38. Fire, et al., Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*, *Nature* 391 (6669) (1998) 806–811.
39. Timmons, et al., Inducible systemic RNA silencing in *Caenorhabditis elegans*, *Mol. Biol. Cell* 14 (7) (2003) 2972–2983.
40. Timmons, et al., Ingestion of bacterial expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*, *Gene* 263 (1–2) (2001) 103–112.
41. Tabara, et al., RNAi in *C. elegans*: soaking in the genome sequence, *Science* 282 (5388) (1998) 430–431.
42. Tavernakis, et al., Heritable and inducible genetics interference by double-stranded RNA encoded by transgenes, *Nat. Genet.* 24 (2) (2000) 180–183.
43. Winston, et al., Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1, *Science* 295 (5564) (2002) 2456–2459.
44. Hunter, et al., Systemic RNAi in *Caenorhabditis elegans*, *Cold Spring Harb. Symp. Quant. Biol.* 71 (2006) 95–100.
45. Simmer, et al., Loss of putative RNA-directed RNA polymerase RRF-3 makes *C. elegans* hypersensitive to RNAi, *Curr. Biol.* 12 (15) (2002) 1317–1319.
46. Kennedy, et al., A conserved siRNA-degrading RNase negatively regulates RNA interference in *C. elegans*, *Nature* 427 (6975) (2004) 645–649.
47. Calixto, et al., Enhanced neuronal RNAi in *C. elegans* using SID-1, *Nat. Methods* 7 (7) (2010) 554–559.
48. Durieux, et al., The cell-non-autonomous nature of electron transport chain-mediated longevity, *Cell* 144 (1) (2011) 79–91.
49. Kamath, et al., Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*, *Genome Biol.* 2 (1) (2001) (research 0002.1-0002.10).
50. Kamath, et al., Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi, *Nature* 421 (6920) (2003) 231–237.
51. Wang, Barr, RNA interference in *Caenorhabditis elegans*, *Methods Enzymol.* 392 (2005) 36–55.
52. Rual, et al., Toward improving *Caenorhabditis elegans* phenome mapping with an ORFeome-based RNAi library, *Genome Res.* 14 (10B) (2004) 2162–2168.
53. Eliscovich, et al., mRNA on the move: the road to its biological destiny, *J. Biol. Chem.* 288 (28) (2013) 20361–20368.
54. Barrett, et al., Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements, *Cell. Mol. Life Sci.* 69 (21) (2012) 3613–3634.
55. Maine, Eleanor, Studying gene function in *Caenorhabditis elegans* using RNA-mediated interference, *Brief. Funct. Genomic Proteomic* 7 (3) (2008) 184–194.
56. Calixto, et al., Conditional gene expression and RNAi using MEC-8-dependent splicing in

- C.elegans*, Nat. Methods 7 (5) (2010) 407–411.
57. Colaiácovo, et al., A targeted RNAi screen for genes involved in chromosome morphogenesis and nuclear organization in the *Caenorhabditis elegans* germline, Genetics 162 (1) (2002) 113–128.
 58. Reinke, et al., A global profile of germline gene expression in *C. elegans*, Mol. Cell 6 (2000) 605–616.
 59. T.J. van Ham, K.L. Thijssen, R. Breitling, R.M. Hofstra, R.H. Plasterk, E.A. Nollen, *C. elegans* model identifies genetic modifiers of alpha-synuclein inclusion formation during aging, PLoS Genet. 4 (2008) e1000027.
 60. van der Goot, et al., Delaying aging and the aging-associated decline in protein homeostasis by inhibition of tryptophan degradation, Proc. Natl. Acad. Sci. U. S. A. 109 (37) (2012) 14912–14917.
 61. Simmer, et al., Genome-wide RNAi of *C. elegans* using the hypersensitive rrf-3 strain reveals novel gene functions, PLoS Biol. 1 (1) (2003) E12.
 62. T.J. van Ham, R. Breitling, M.A. Swertz, E.A. Nollen, Neurodegenerative diseases: lessons from genome-wide screens in small model organisms, EMBO Mol. Med. 1 (2009) 360–370.
 63. C. Soto, Unfolding the role of protein misfolding in neurodegenerative diseases, Nat. Rev. Neurosci. 4 (2003) 49–60.
 64. F. Chiti, C.M. Dobson, Protein misfolding, functional amyloid, and human disease, Annu. Rev. Biochem. 75 (2006) 333–366.
 65. B. Bolognesi, J.R. Kumita, T.P. Barros, E.K. Esbjorner, L.M. Luheshi, D.C. Crowther, M.R. Wilson, C.M. Dobson, G. Favrin, J.J. Yerbury, ANS binding reveals common features of cytotoxic amyloid species, ACS Chem. Biol. 20 (2010) 735–740.
 66. M. Arrasate, S. Mitra, E.S. Schweitzer, M.R. Segal, S. Finkbeiner, Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death, Nature 404 (2004) 805–810.
 67. C. Bessou, J.B. Giuglia, C.J. Franks, L. Holden-Dye, L. Ségalat, Mutations in the *Caenorhabditis elegans* dystrophin-like gene dys-1 lead to hyperactivity and suggest a link with cholinergic transmission, Neurogenetics 2 (1998) 61–72.
 68. J.F. Morley, H.R. Brignull, J.J. Weyers, R.I. Morimoto, The threshold for polyglutamine expansion protein aggregation and cellular toxicity is dynamic and influenced by aging in *Caenorhabditis elegans*, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 10417–10422.
 69. S. Hamamichi, R.N. Rivas, A.L. Knight, S. Cao, K.A. Caldwell, G.A. Caldwell, Hypothesis-based RNAi screening identifies neuroprotective genes in a Parkinson's disease model, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 728–733.
 70. T. Kuwahara, A. Koyama, S. Koyama, S. Yoshina, C.H. Ren, T. Kato, S. Mitani, T. Iwatsubo, A systematic RNAi screen reveals involvement of endocytic pathway in neuronal dysfunction in alpha-synuclein transgenic *C. elegans*, Hum. Mol. Genet. 17 (2008) 2997–3009.
 71. C.D. Link, Expression of human beta-amyloid peptide in transgenic *Caenorhabditis elegans*, Proc. Natl. Acad. Sci. U. S. A. 92 (1995) 9368–9372.
 72. N. Wittenburg, S. Eimer, B. Lakowski, S. Roëhrig, C. Rudolph, R. Baumeister, Presenilin is required for proper morphology and function of neurons in *C. elegans*, Nature 406 (2000) 306–309.
 73. E. Cohen, J. Bieschke, R.M. Perciavalle, J.W. Kelly, A. Dillin, Opposing activities protect against age-onset proteotoxicity, Science 313 (2006) 1604–1610.
 74. E. Cohen, J.F. Paulsson, P. Blinder, T. Burstyn-Cohen, D. Du, G. Estepa, A. Adame, H.M. Pham, M. Holzenberger, J.W. Kelly, E. Masliah, A. Dillin, Reduced IGF-1 signaling delays age-associated proteotoxicity in mice, Cell 139 (2009) 1157–1169.
 75. B.C. Kraemer, J.K. Burgess, J.H. Chen, J.H. Thomas, G.D. Schellenberg, Molecular pathways that influence human tau-induced pathology in *Caenorhabditis elegans*, Hum. Mol. Genet. 15 (2006) 1483–1496.
 76. T.J. van Ham, M.A. Holmberg, A.T. van der Goot, E. Teuling, M. Garcia-Arencibia, H.E. Kim, D. Du, K.L. Thijssen, M. Wiersma, R. Burggraaff, P. van Bergeijk, J. van Rheenen, G.J. van Veluw, R.M. Hofstra, D.C. Rubinsztein, E.A. Nollen, Identification of MOAG-4/SERF as a regulator of age-related proteotoxicity, Cell 140 (2010) 601–612.
 77. F.X. Lejeune, L. Mesrob, F. Parmentier, C. Bicep, R.P. Vazquez-Manrique, J. Alex Parker, J.P. Vert, C. Tourette, C. Neri, Large-scale functional RNAi screen in *C. elegans* identifies genes that regulate the dysfunction of mutant polyglutamine neurons, BMC Genomics 13 (2012) 91.
 78. M. Komatsu, Q.J. Wang, G.R. Holstein, V.L. Friedrich Jr., J. Iwata, E. Kominami, B.T. Chait, K. Tanaka, Z. Yue, Essential role for autophagy protein Atg7 in the maintenance of axonal homeostasis and the prevention of axonal degeneration, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 14489–14494.

79. E.S. Blum, A.R. Schwendeman, S. Shaham, PolyQ disease: misfiring of a developmental cell death program? *Trends Cell Biol.* 23 (2013) 168–174.
80. P.W. Faber, J.R. Alter, M.E. MacDonald, A.C. Hart, Polyglutamine-mediated dysfunction and apoptotic death of a *Caenorhabditis elegans* sensory neuron, *Proc. Natl. Acad. Sci. U. S. A.* 96 (1999) 179–184.
81. J.A. Parker, J.B. Connolly, C. Wellington, M. Hayden, J. Dausset, C. Neri, Expanded polyglutamines in *Caenorhabditis elegans* cause axonal abnormalities and severe dysfunction of PLM mechanosensory neurons without cell death, *Proc. Natl. Acad. Sci. U. S. A.* 98 (2001) 13318–13323.
82. H.R. Brignull, F.E. Moore, S.J. Tang, R.I. Morimoto, Polyglutamine proteins at the pathogenic threshold display neuron-specific aggregation in a pan-neuronal *Caenorhabditis elegans* model, *J. Neurosci.* 26 (2006) 7597–7606.
83. E.A. Nollen, S.M. Garcia, H.G. van Haften, S. Kim, A. Chavez, R.I. Morimoto, R.H. Plasterk, Genome-wide RNA interference screen identifies previously undescribed regulators of polyglutamine aggregation, *Proc. Natl. Acad. Sci. U. S. A.* 101 (2004) 6403–6408.
84. M.C. Silva, S. Fox, M. Beam, H. Thakkar, M.D. Amaral, R.I. Morimoto, A genetic screening strategy identifies novel regulators of the proteostasis network, *PLoS Genet.* 7 (2011) e1002438.
85. S.F. Falsone, et al, SERF Protein Is a Direct Modifier of Amyloid Fiber Assembly, *Cell.* 2 (2012) 358–371.
86. L.M. de Lau, M.M. Breteler, Epidemiology of Parkinson's disease, *Lancet Neurol.* 5 (2006) 525–535.
87. N.N. Inamdar, D.K. Arulmozhi, A. Tandon, S.L. Bodhankar, Parkinson's disease: genetics and beyond, *Curr. Neuropharmacol.* 5 (2007) 99–113.
88. S. Przedborski, W. Dauer, Parkinson's disease: mechanisms and models, *Neuron* 39 (2003) 889–909.
89. K.K. Chung, Y. Zhang, K.L. Lim, Y. Tanaka, H. Huang, J. Gao, C.A. Ross, V.L. Dawson, T.M. Dawson, Parkin ubiquitinates the α -synuclein-interacting protein, synphilin-1: implications for Lewy-body formation in Parkinson disease, *Nat. Med.* 7 (2001) 1144–1150.
90. A. Abeliovich, Y. Schmitz, I. Fariñas, D. Choi-Lundberg, W.H. Ho, P.E. Castillo, N. Shinsky, J.M. Verdugo, M. Armanini, A. Ryan, M. Hynes, H. Phillips, D. Sulzer, A. Rosenthal, Mice lacking alpha-synuclein display functional deficits in the nigrostriatal dopamine system, *Neuron* 25 (2000) 239–252.
91. C. Lavedan, The synuclein family, *Genome Res.* 8 (1998) 871–880.
92. D.F. Clayton, J.M. George, The synucleins: a family of proteins involved in synaptic function, plasticity, neurodegeneration and disease, *Trends Neurosci.* 21 (1998) 249–254.
93. M.H. Polymeropoulos, C. Lavedan, E. Leroy, S.E. Ide, A. Dehejia, A. Dutra, B. Pike, H. Root, J. Rubenstein, R. Boyer, E.S. Stenroos, S. Chandrasekharappa, A. Athanassiadou, T. Papapetropoulos, W.G. Johnson, A.M. Lazzarini, R.C. Duvoisin, G. DiIorio, L.I. Golbe, R.L. Nussbaum, Mutation in the alpha-synuclein gene identified in families with Parkinson's disease, *Science* 276 (1997) 2045–2047.
94. R. Krüger, W. Kuhn, T. Müller, D. Woitalla, M. Graeber, S. Kösel, H. Przuntek, J.T. Eppelen, L. Schöls, O. Riess, Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease, *Nat. Genet.* 18 (1998) 106–108.
95. J.J. Zarranz, J. Alegre, J.C. Gómez-Esteban, E. Lezcano, R. Ros, I. Ampuero, L. Vidal, J. Hoenicka, O. Rodríguez, B. Atarés, V. Llorens, E. Gomez Tortosa, T. del Ser, D.G. Muñoz, J.G. de Yebenes, The new mutation, E46K, of alpha-synuclein causes Parkinson and Lewy body dementia, *Ann. Neurol.* 55 (2004) 164–173.
96. M.C. Chartier-Harlin, J. Kachergus, C. Roumier, V. Mouroux, X. Douay, S. Lincoln, C. Levecque, L. Larvor, J. Andrieux, M. Hulihan, N. Waucquier, L. Defebvre, P. Amouyel, M. Farrer, A. Destée, Alpha-synuclein locus duplication as a cause of familial Parkinson's disease, *Lancet* 364 (2004) 1167–1169.
97. A.B. Singleton, M. Farrer, J. Johnson, A. Singleton, S. Hague, J. Kachergus, M. Hulihan, T. Peuralinna, A. Dutra, R. Nussbaum, S. Lincoln, A. Crawley, M. Hanson, D. Maraganore, C. Adler, M.R. Cookson, M. Muentner, M. Baptista, D. Miller, J. Blancato, J. Hardy, K. Gwinn-Hardy, Alpha-synuclein locus triplication causes Parkinson's disease, *Science* 302 (2003) 841.
98. Fung, et al., Genome-wide genotyping in Parkinson's disease and neurologically normal controls: first stage analysis and public release of data, *Lancet Neurol.* 5 (2006) 911–916.
99. B.M. Collins, A.J. McCoy, H.M. Kent, P.R. Evans, D.J. Owen, Molecular architecture and functional model of the endocytic AP2 complex, *Cell* 109

- (2002) 523–535.
100. K. Bettens, K. Sleegers, C. Van Broeckhoven, Genetic insights in Alzheimer's disease, *Lancet Neurol.* 12 (2013) 92–104.
 101. C. Ballard, S. Gauthier, A. Corbett, C. Brayne, D. Aarsland, E. Jones, Alzheimer's disease, *Lancet* 377 (2011) 1019–1031.
 102. A.J. Myers, M. Kaleem, L. Marlowe, A.M. Pittman, A.J. Lees, H.C. Fung, J. Duckworth, D. Leung, A. Gibson, C.M. Morris, R. de Silva, J. Hardy, The H1c haplotype at the MAPT locus is associated with Alzheimer's disease, *Hum. Mol. Genet.* 14 (2005) 2399–23404.
 103. T.M. Caffrey, R. Wade-Martins, Functional MAPT haplotypes: bridging the gap between genotype and neuropathology, *Neurobiol. Dis.* 27 (2007) 1–10.
 104. G. McGoll, B.R. Roberts, A.P. Gunn, K.A. Perez, D.J. Tew, C.L. Masters, K.J. Barnham, R.A. Cherny, A.I. Bush, The *Caenorhabditis elegans* A β 1–42 model of Alzheimer disease predominantly expresses A β 3–42, *J. Biol. Chem.* 284 (2009) 22697–22702.
 105. C.D. Link, A. Taft, V. Kapulkin, K. Duke, S. Kim, Q. Fei, D.E. Wood, B.G. Sahagan, Gene expression analysis in a transgenic *Caenorhabditis elegans* Alzheimer's disease model, *Neurobiol. Aging* 24 (2003) 397–413.
 106. Y. Wu, Z. Wu, P. Butko, Y. Christen, M.P. Lambert, W.L. Klein, C.D. Link, Y. Luo, Amyloid-beta-induced pathological behaviors are suppressed by Ginkgo biloba extract EGb 761 and ginkgolides in transgenic *Caenorhabditis elegans*, *J. Neurosci.* 26 (2006) 13102–13113.
 107. G. McColl, B.R. Roberts, T.L. Pukala, V.B. Kenche, C.M. Roberts, C.D. Link, T.M. Ryan, C.L. Masters, K.J. Barnham, A.I. Bush, R.A. Cherny, Utility of an improved model of amyloid-beta (A β 1–42) toxicity in *Caenorhabditis elegans* for drug screening for Alzheimer's disease, *Mol. Neurodegener.* 7 (2012) 57.
 108. T. Miyasaka, Z. Ding, K. Gengyo-Ando, M. Oue, H. Yamaguchi, S. Mitani, Y. Ihara, Progressive neurodegeneration in *C. elegans* model of tauopathy, *Neurobiol. Dis.* 20 (2005) 372–383.
 109. R. Brandt, A. Gergou, I. Wacker, T. Fath, H. Hutter, A *Caenorhabditis elegans* model of tau hyperphosphorylation: induction of developmental defects by transgenic overexpression of Alzheimer's disease-like modified tau, *Neurobiol. Aging* 30 (2009) 22–33.
 110. H.Y. Wang, W.W. Li, N.J. Benedetti, D.H.S. Lee, Alpha(7) nicotinic acetylcholine receptors mediate beta-amyloid peptide-induced tau protein phosphorylation, *J. Biol. Chem.* 278 (2003) 31547–31553.
 111. T. Kaletta, M.O. Hengartner, Finding function in novel targets: *C. elegans* as a model organism, *Nat. Rev. Drug Discov.* 5 (2006) 387–398.
 112. S. Treusch, S. Hamamichi, J.L. Goodman, K.E. Matlack, C.Y. Chung, V. Baru, J.M. Shulman, A. Parrado, B.J. Bevis, J.S. Valastyan, H. Han, M. Lindhagen-Persson, E.M. Reiman, D.A. Evans, D.A. Bennett, A. Olofsson, P.L. DeJager, R.E. Tanzi, K.A. Caldwell, G.A. Caldwell, S. Lindquist, Functional links between A β toxicity, endocytic trafficking, and Alzheimer's disease risk factors in yeast, *Science* 334 (2011) 1241–1245.
 113. J.M. Shulman, P. Chipendo, L.B. Chibnik, C. Aubin, D. Tran, B.T. Keenan, P.L. Kramer, J.A. Schneider, D.A. Bennett, M.B. Feany, P.L. De Jager, Functional screening of Alzheimer pathology genome-wide association signals in *Drosophila*, *Am. J. Hum. Genet.* 88 (2011) 232–238.





CHAPTER IV

Functional sequestration of MOAG-2/LIR-3 by polyglutamine expansion proteins

Olga Sin^{1,2}, Tristan de Jong¹, Michelle Kudron³, Renée Seinstra¹, Alejandro Mata-Cabana¹, Céline Martineau¹, Roméo Willinge Prins¹, Hai Hui Wang¹, Wytse Hogewerf¹, Valerie Reinke³, Victor Guryev¹, Ellen Nollen¹

¹ European Research Institute for the Biology of Aging, Groningen University of Groningen, University Medical Centre Groningen, The Netherlands

² Graduate Program in Areas of Basic and Applied Biology, University of Porto, Porto, Portugal

³ Department of Genetics, Yale University School of Medicine, New Haven, USA.

Manuscript in preparation

Abstract

Aging-related protein aggregation is one of the hallmarks of neurodegenerative disorders such as Alzheimer's, Parkinson's and polyglutamine diseases. The cellular processes that drive protein aggregation in these diseases have remained largely unknown. Using a forward genetic screen in a *C. elegans* model for polyglutamine aggregation, we here identified *lin-26*-related gene 3 (*lir-3*) as modifier of protein aggregation (*moag-2*) that promotes protein aggregation. A mutation or a deletion in *moag-2/lir-3* reduced protein aggregation. *moag-2/lir-3* encodes a protein of unknown function with a putative nuclear localization domain and two domains that are homologous to nucleotide-binding C2H2 zinc finger domains. By combining chromatin immunoprecipitation sequencing and RNA sequencing, we found that MOAG-2/LIR-3 is preferentially associated with promoter regions of RNA Polymerase III-regulated non-coding RNAs and regulates their expression in wild type N2 nematodes. This regulation is lost in worms that express proteins with expanded polyglutamine tracts. These results suggest that, by driving polyglutamine aggregation, MOAG-2/LIR-3 can no longer execute its normal function thus resulting in a transcriptional change of small non-coding RNAs. Together, our results show that aggregation-prone disease proteins can turn benign cellular proteins into aggregation-promoting factors, at the expense of their biological function. The functional switch can be explored as a therapeutic target to interfere with aggregation events in neurodegenerative diseases.

Introduction

Neurodegenerative diseases like Alzheimer's, Parkinson's or polyglutamine diseases are a major health concern and demand a better understanding of the disease pathogenesis in order to make disease-modifying treatments available. One core pathological hallmark of neurodegenerative diseases is the presence of protein aggregates in different brain areas of affected patients [1]. These insoluble, macromolecular structures are particularly enriched in aggregation-prone proteins which, by exposing regions of their amino acid sequence, begin to self-associate or aggregate with other proteins of the cell, thereby hampering normal cellular function [2-4]. Depending on whether these proteins remain soluble and misfolded or acquire amyloidogenic properties they will be localized to different cellular compartments where they can either be degraded and recycled by the ubiquitin-proteasome machinery or form insoluble aggregates [5, 6]. It is not resolved whether these protein aggregates are causing disease. The current view is that soluble oligomeric or fibrillar precursors to these aggregates are cytotoxic and that aggregation is a protective measure to sequester these harmful species [7-10].

The cellular factors that drive protein aggregation are poorly understood. Direct aggregation-promoting factors have been identified [11-14]. SH3GL3 (or endophilin-3) is a SH3-domain GRB2-like 3 protein that was suggested to interact directly with the polyglutamine stretch within exon 1 of the *huntingtin* gene to promote aggregation [14]. This interaction was proposed to be dependent on the polyglutamine length of mutant *huntingtin* or its aggregated form [14, 15]. MOAG-4/SERF was identified as a direct modifier of polyglutamine, alpha-synuclein and amyloid-beta-induced aggregation [11, 12]. MOAG-4/SERF is a predominantly disordered protein of unknown function that was shown to drive aggregation by transient binding to aggregation-prone proteins at the early stages of the aggregation pathway [11, 12]. Cell non-autonomous regulation of protein aggregation has been described with the example of *unc-30*, a homeodomain transcription factor involved in the synthesis of the inhibitory neurotransmitter GABA [16]. Mutation of *unc-30* impairs neuronal signaling at the neuromuscular junction and enhances protein aggregation in a *C. elegans* model for polyglutamine diseases [16].

A wide variety of genetic screens have been performed to find genes that regulate aggregation and aggregation-associated toxicity in animal models for neurodegenerative diseases [12, 17-22]. Here, we describe a forward genetic screen where we sought for genes that normally drive protein aggregation in a *C. elegans* model for polyglutamine diseases. We identified *moag-2/lir-3* as a modifier of aggregation that, when mutated, reduced aggregation up to 51% in a polyglutamine model. We demonstrate that MOAG-2/LIR-3 is a transcriptional regulator that promotes the transcription of small non-coding RNAs (ncRNA), including small nucleolar RNAs (snoRNAs) and transfer RNAs (tRNAs), in wild type nematodes. We propose that MOAG-2/LIR-3 shifts its role in the presence of polyglutamine from a transcriptional regulator of ncRNAs to an aggregation-promoting factor.

Results

Inactivation of MOAG-2/LIR-3 reduces polyglutamine aggregation

To identify genes that drive protein aggregation, we performed a forward genetic screen in a *C. elegans* model that expresses an aggregation-prone polyglutamine stretch of 40 residues fused to YFP (Q40-YFP) in the body-wall muscle cells [12]. We screened for mutants that reduced polyglutamine aggregation, which we named modifiers of aggregation (*moag*) ([12] and unpublished data). *moag-2(pk2183)* (hereafter designated simply as "*lir-3(pk2183)*") presented a 51% reduction in the number of aggregates relative to the wild type Q40 animals in their fourth larval stage (Figure 1A, B). SNP-mapping and genome sequencing allowed us to fine-map the causal mutation and revealed six genes as putative candidates for *moag-2* (Figure S1B). For one of these, which mapped to *lir-3* (*lin-26* related; sequence: F37H8.1; accession number: NC_003280), the causative mutation was in the start codon, replacing the first methionine with an isoleucine (Met1Ile) (Figure 1C, S1B). *lir-3* encodes a LIN-26-like zinc finger protein of unknown function (<http://wormbase.org>, March 2014). It is predicted to have two zinc finger domains of the C2H2 type (residues 191-214 and 224-247) at the carboxyl terminus (Figure S1C) and a nuclear localization signal spanning amino acid residues 132 to 141 (<http://nls-mapper.iab.keio.ac.jp/>, March 2014). LIR-3 shares two non-canonical C2H2 zinc finger motifs with LIN-26 (31 to 35% identical), LIR-1 (20 to 31% identical) and LIR-2 (25% identical) [23].

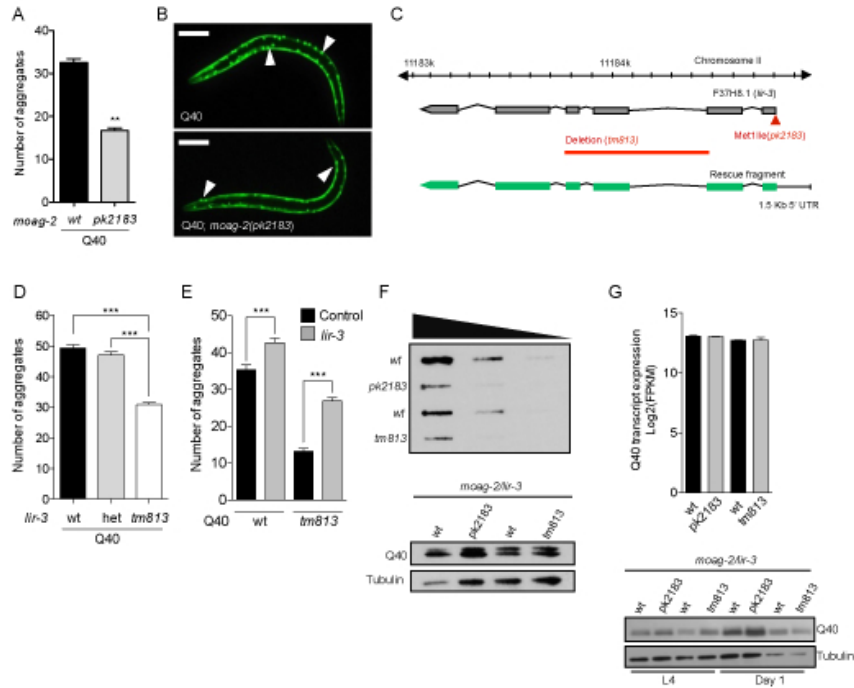


Figure 1. Identification of *moag-2/lir-3* as an aggregation-promoting factor. (A) Number of aggregates in Q40 animals and Q40;*moag-2(pk2183)* animals. (B) Representative images of Q40 animals and Q40;*moag-2(pk2183)* animals. Scale bar, 75 μ m (C) Chromosomal location of *lir-3* (F37H8.1, Chromosome II, reverse strand of assembly; <http://www.wormbase.org>, WS248), showing the point mutation in the start codon (red arrowhead), the partial deletion (red bar) and the rescue fragment (green). (D) Number of aggregates in Q40 animals with wild type, heterozygous and homozygous deletion for the *lir-3(tm813)* allele. (E) Number of aggregates in Q40 animals and Q40;*lir-3(tm813)* animals, with and without transgenic overexpression of an injected *lir-3* construct, using its endogenous promoter. (F) Filter retardation assay with five-fold serial dilutions of crude protein extract from Q40 animals and Q40;*moag-2/lir-3* mutant animals. See also Figure S1D. Q40-YFP and α -tubulin expression were included as controls. (G) Q40-YFP transcript expression detected by RNA seq in L4 animals and protein expression detected in urea-treated L4 and day one adult animals. In all panels, aggregate counting, representative images and filter retardation assay were performed at the L4 stage and represented is the average of three biological replicates. Data are represented as mean \pm SEM and significance was calculated using one-tailed unpaired t-test. ** $p < 0.01$; *** $p < 0.001$

To assess whether *moag-2* was *lir-3*, we generated a *lir-3(tm813)* deletion mutant (hereafter designated simply as "*lir-3(tm813)*") and crossed it with the polyglutamine animals. This strain has a 795 bp deletion that spans residues 276 to 1070 of the F37H8.1 sequence that causes a premature stop codon (Figure 1C, S1C). The partial deletion of *lir-3* caused a 35% reduction of aggregates relative to the wild type animals (Figure 1D). Animals heterozygous for the *lir-3* deletion allele exhibited a similar number of aggregates as wild type, supporting that the reduction of aggregation was recessive and due to the loss of function of *lir-3* (Figure 1D). We next asked whether overexpression

of *lir-3* could restore the aggregation phenotype. We injected a rescue construct with full-length *lir-3* including a 1.5 kb sequence upstream of the start codon to include its endogenous promoter as well as 330 bp downstream to include the 3'UTR (Figure 1C). The expression of the *lir-3* rescue fragment in Q40;*lir-3(tm813)* animals was able to restore the aggregation phenotype by 2-fold ($p < 0.05$), therefore confirming *lir-3* as the gene responsible for driving aggregation in our polyglutamine model (Figure 1E). In summary, these results demonstrate that *lir-3* is the *moag-2* gene.

One property of aggregates in the brains of neurodegenerative disease patients, which is also captured by the *C. elegans* polyglutamine model, is that they are typically resistant to strong detergents, such as sodium dodecyl sulfate (SDS) [24-26]. To establish whether MOAG-2/LIR-3 promoted aggregation of SDS-insoluble polyglutamine aggregates, we performed a filter retardation assay on lysates of wild type and *moag-2/lir-3* mutant polyglutamine animals [27, 28]. This assay enables the detection of SDS-insoluble protein aggregates while smaller, soluble species are not captured. Both Q40;*moag-2/lir-3(pk2183)* and Q40;*moag-2/lir-3(tm813)* mutants presented less SDS-insoluble aggregates than their corresponding Q40 controls (Figure 1F). The reduction of SDS-insoluble aggregates was more pronounced in the point mutant (43%; $p = 0.491$) than in the deletion mutant (26%; $p = 0.1828$) (Figure S1D). Mutation or partial deletion of *moag-2/lir-3* did not detectably reduce the transcription or the protein expression level of Q40-YFP, indicating that *moag-2/lir-3* does not reduce aggregation by reducing expression level of the Q40-YFP protein (Figure 1F, G). Together, these results indicate that *moag-2/lir-3* drives the formation of SDS-insoluble aggregates.

MOAG-2/LIR-3 is C2H2-domain protein associated with RNA Polymerase III promoters

Having established that mutation of *moag-2/lir-3* reduces polyglutamine aggregation, we next investigated the endogenous function of the protein. C2H2 zinc finger domains are predominantly associated with DNA-binding transcription factors but may also have other functions such as mediating protein-protein interactions or binding to RNA [29-31]. Bioinformatic analysis combined with manual curation predicted MOAG-2/LIR-3 to be a

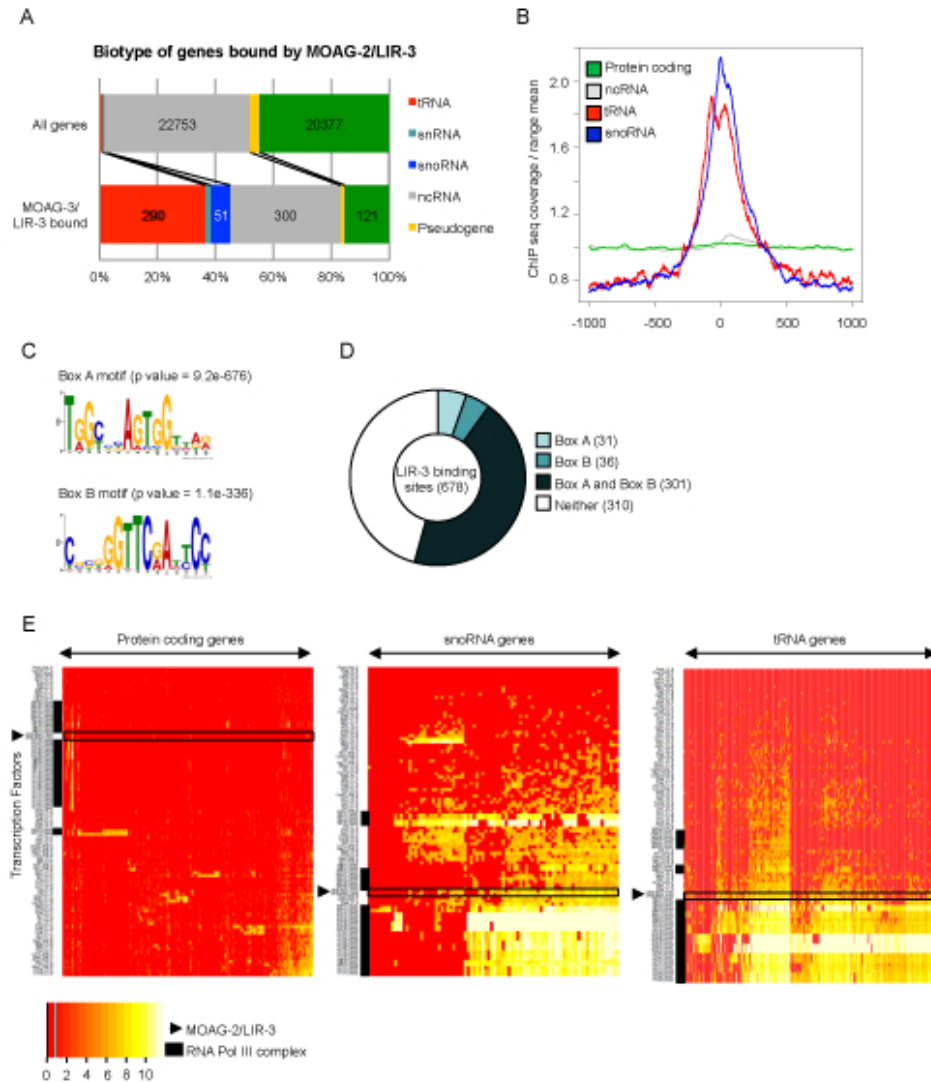


Figure 2. MOAG-2/LIR-3 preferentially binds to promoters of small ncRNAs. (A) Comparison between the composition of the gene biotypes bound by MOAG-2/LIR-3 (1kb around TSS) and those distributed genome-wide. See also Figure S2A. (B) MOAG-2/LIR-3 binding sites within -1000 and +1000 bp from TSS for protein-coding, ncRNA, tRNA and snoRNA genes. (C) Enriched consensus DNA motifs for MOAG-2/LIR-3 with p value. (D) Number of MOAG-2/LIR-3 binding sites containing Box A and Box B. (E) Heat map showing the binding of different transcription factors to promoters of *C. elegans* protein-coding, snoRNA and tRNA genes. The hierarchical clustering was generated using the average linkage cluster method with a binary metric distance.

transcription factor [32, 33]. To explore this possibility, we took advantage of chromatin immunoprecipitation followed by deep sequencing (ChIP seq) as previously described, using L4-staged animals that express an integrated construct of *lir-3* fused to a GFP tag [34, 35]. This analysis yielded a total of 678 unique MOAG-2/LIR-3 binding sites, of which 404 overlapped with 813 *C. elegans* genes. Further analyses of these genes revealed that MOAG-2/LIR-3 binding was enriched in the transcription start sites (TSS) of tRNA genes (35.7%, $p < 0.001$), snoRNA genes (6.3%, $p < 0.001$), rRNA genes (2.5%, $p < 0.001$) and snRNA genes (2.2%, $p < 0.002$) (Figure 2A, B, S2A). While MOAG-2/LIR-3 was also found in the vicinity of protein-coding and other ncRNA genes, this binding was not significantly enriched (Figure 2A, B).

We then asked whether the binding sites were enriched in any consensus sequence motif that could be recognized by MOAG-2/LIR-3. From the 678 binding sites initially identified in our ChIP seq, more than half of the sites contained Box A and Box B sequence motifs, 301 of which contained both motifs (Figure 2C, D). Box A and Box B constitute the canonical type 2 promoter site recognized by the RNA Polymerase (Pol) III complex [36, 37]. Pol III is responsible for the transcription of structural or catalytic small nuclear RNAs (snRNAs), of tRNAs and of snoRNAs, which mediate chemical modifications of other RNA molecules [37-40].

These findings led us to hypothesize that MOAG-2/LIR-3 could bind to the same target promoters as Pol III. Several *C. elegans* transcription factors have been shown to bind in the proximity of non-coding genes, including PHA-4, PQM-1 and GEL-11 [41]. This prompted us to ask whether the association of MOAG-2/LIR-3 with the RNA Pol III complex resembled the binding of these transcription factors to the promoters of non-coding genes. To answer this question, we collected publicly available ChIP seq data for known transcription factors of *C. elegans* (<http://www.modencode.org>; September 2014) and analyzed their binding to the promoters of protein-coding genes, snoRNA and tRNA genes (Figure 2E). While there is little association of MOAG-2/LIR-3 with the promoters of protein-coding genes, there is a strong similarity of MOAG-2/LIR-3 binding profile with a group of factors clustered together with representative components of the RNA Pol III complex, including

Pol III, the TATA binding protein (TBP-1), two subunits of the transcription factor for Pol III C (TFC-1 and TFC-4) and the nuclear pore proteins NPP-3 and NPP-13, which were recently shown to associate with the RNA Pol III complex to regulate tRNA and snoRNA splicing [36]. The binding of these factors is very abundant in the promoters of snoRNAs as well as tRNAs (Figure 2E, arrowhead), in contrast to the majority of the other transcription factors. Together, we demonstrated that MOAG-2/LIR-3 binds to the promoters of small ncRNA genes suggesting that MOAG-2/LIR-3 is associated with Pol III transcription.

MOAG-2/LIR-3 is a positive regulator of the RNA Pol III-mediated transcription of small ncRNAs

Because MOAG-2/LIR-3 bound to the promoters of small ncRNA genes, we next asked what consequence this binding would have for the transcription of the Pol III downstream targets. We therefore compared the RNA expression in wild type animals to *moag-2/lir-3* mutant animals by transcriptome profiling. In line with the absence of MOAG-2/LIR-3 at Pol II promoter sites, we did not find any differentially expressed protein-coding genes in the mutants compared to the wild-type N2 animals, thereby excluding MOAG-2/LIR-3 as a transcriptional regulator of protein-coding genes (Figure 3C, S3A). Mutations in *moag-2/lir-3*, however, did result in the downregulation of snRNAs ($p < 0.001$), snoRNAs ($p < 0.001$) and tRNAs ($p < 0.001$) in both mutants, demonstrating that MOAG-2/LIR-3 is required for Pol III-mediated transcription of these small ncRNAs (Figure 3C, S3A). We next asked where MOAG-2/LIR-3 was positioned relative to RNA Pol III complex, by comparing the positions of the ChIP seq signals of MOAG-2/LIR-3 relative to the different components of the RNA Pol III complex. For both the tRNA and snoRNA genes, all factors localized to the Box A and Box B containing promoter region, consistent with previous reports (Figure 3A, B) [36]. MOAG-2/LIR-3 was positioned to the same sites (Figure 3A, B). Together, these results indicate that MOAG-2/LIR-3 functions as a positive regulator of the Pol III-mediated transcription of small ncRNAs in *C. elegans*.

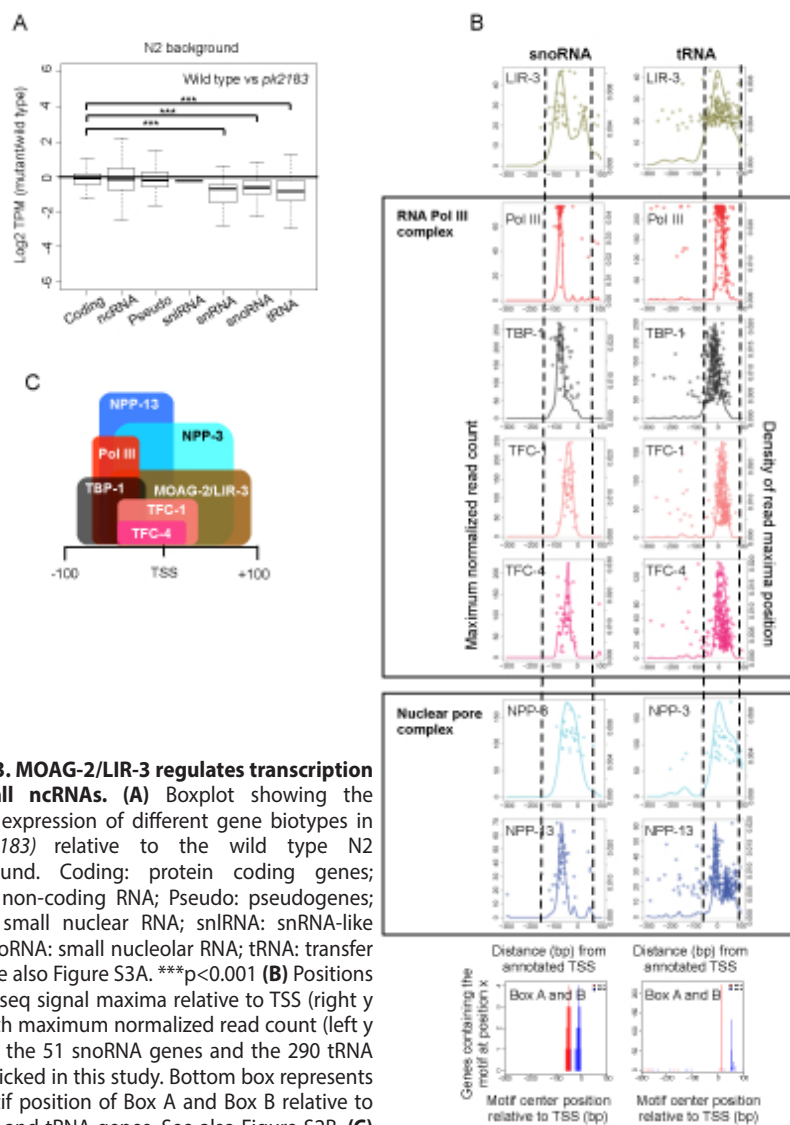


Figure 3. MOAG-2/LIR-3 regulates transcription of small ncRNAs. (A) Boxplot showing the relative expression of different gene biotypes in *lir-3(pk2183)* relative to the wild type N2 background. Coding: protein coding genes; ncRNA: non-coding RNA; Pseudo: pseudogenes; snRNA: small nuclear RNA; snRNA-like RNA; snoRNA: small nucleolar RNA; tRNA: transfer RNA. See also Figure S3A. *** $p < 0.001$ (B) Positions of ChIP seq signal maxima relative to TSS (right y axis) with maximum normalized read count (left y axis) for the 51 snoRNA genes and the 290 tRNA genes picked in this study. Bottom box represents the motif position of Box A and Box B relative to snoRNA and tRNA genes. See also Figure S2B. (C) Diagram showing the positions of the RNA Pol III factors and MOAG-2/LIR-3 estimated from data presented in panel (B).

Regulation of protein aggregation by MOAG-2/LIR-3 is independent of its role as a transcriptional regulator

Next, we asked whether MOAG-2/LIR-3 regulated protein aggregation via the RNA Pol III-mediated transcription of the small ncRNAs. We therefore knocked down by RNAi, one-by-one, the individual components of the RNA Pol III complex in both the wild type and the *moag-2/lir-3(pk2183)* mutants. To confirm RNAi knockdown, we also looked for RNAi-associated phenotypes other than aggregation. If Pol III-mediated transcription would play a role, this would result in a reduction in the amount of aggregates in the wild type Q40 animals but not in the *moag-2/lir-3* mutants. As a control, knockdown of Pol II decreased the amount of aggregates in both the wild type and the mutant animals, indicating that a reduction in the expression of coding genes has an effect independent of *moag-2/lir-3* (Figure 4A). Knockdown of Pol III, TBP-1, TFC-1 and TFC4 did not alter aggregation in the Q40 nor in the Q40;*moag-2/lir-3* mutant strains, indicating that reduction of Pol III-mediated transcription in the absence of *moag-2/lir-3* is not responsible for the reduction of aggregation (Figure 4A). Reduction of *npp-13* or of tRNA processing enzymes also did not alter the aggregation phenotype, indicating that neither small RNA processing nor the availability of mature tRNAs plays a role in reduction of aggregation (Figure 4A, B). These results indicate that the regulation of protein aggregation by MOAG-2/LIR-3 is separate from its role in RNA transcription with the RNA Pol III complex.

Polyglutamine expansion proteins suppress transcription of small ncRNAs

We next aimed to address why the role of MOAG-2/LIR-3 in driving protein aggregation was independent of its role as a transcriptional regulator. We therefore first compared the RNA expression profiles of wild type Q40 animals to *moag-2/lir-3* mutant Q40 animals. In contrast to wild type N2 animals, in the presence of polyglutamine there was no longer a change in the relative expression levels of the small ncRNAs (Figure 4C). When we then measured the absolute levels of all RNAs, we found that, in contrast to the protein-coding RNAs, pseudogenes and other ncRNAs, the expression of snRNAs, snoRNAs and tRNAs was already strongly reduced in wild type Q40 animals, indicating that polyglutamine expansion proteins downregulated ncRNA expression

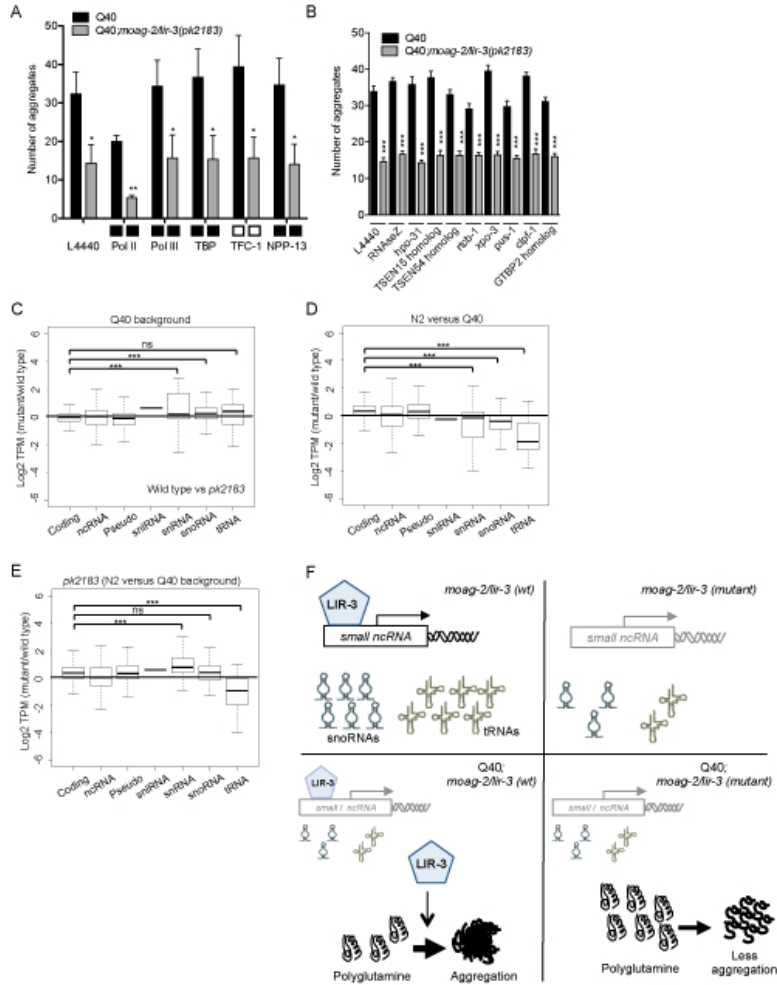


Figure 4. MOAG-2/LIR-3 drives aggregation by converting into an aggregation-promoting factor.

(A) Number of aggregates measured upon RNAi knockdown of individual components of the RNA Pol III complex in Q40 and Q40;*lir-3(pk2183)* animals. Aggregate counting was performed at young adult stage. As an internal quality control for RNAi, squares indicate penetrance (100% [closed] and 0% [open]) of all associated visible RNAi phenotypes other than aggregation. See also Figure S4A. (B) Number of aggregates measured upon RNAi knockdown of tRNA processing enzymes in Q40 and Q40;*lir-3(pk2183)* animals. Aggregate counting was performed at L4 stage. In panels A and B, data are represented as mean \pm SEM and significance was calculated using two-tailed unpaired t-test. (C-E) Boxplot showing the relative expression of different gene biotypes in (C) Q40 and Q40;*lir-3(pk2183)*; in (D) N2 and Q40 wild type animals (Q40 wild type outcrossed from *pk2183*); and in (E) *lir-3(pk2183)* in the N2 and Q40 genetic backgrounds. See also Figure S4C-D. Coding: protein coding genes; ncRNA: non-coding RNA; Pseudo: pseudogenes; snRNA: small nuclear RNA; snRNA-like: snRNA-like RNA; snRNA: small nucleolar RNA; tRNA: transfer RNA. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$, ns is not significant. In all panels, the average of three biological replicates is represented. (F) Working hypothesis for the role of MOAG-2/LIR-3 as an aggregation-promoting factor. In wild type animals, MOAG-2/LIR-3 normally regulates the transcription of small ncRNAs, namely snRNAs, snoRNAs and tRNAs. However, when polyglutamine expansion proteins are expressed, MOAG-2/LIR-3 switches its function from transcriptional regulator to an aggregation-promoting factor, thereby driving polyglutamine aggregation.

($p < 0.001$; Figure 4D). We observed this reduction in expression to a lesser extent in the *moag-2/lir-3* mutant Q40 animals compared to the mutant N2 animals, suggesting that part of the downregulation in the wild type animals was caused by polyglutamine inhibition of MOAG-2/LIR-3 ($p < 0.001$; Figure 4E). Indeed, this finding explained why mutations in *moag-2/lir-3* could no longer reduce the expression in these Q40 animals (Figure 4C, S4C).

Altogether, these data suggest a switch of function of MOAG-2/LIR-3 in presence of polyglutamine expansion proteins from its biological role as a transcriptional regulator to a pathological role as an aggregation promoter at the expense of its own function (Figure 4F).

Discussion

Aggregation-prone disease proteins have been found to cause age-related neurodegenerative diseases but the cellular processes that drive their aggregation and toxicity are still not fully understood. Using a genetic screen in a *C. elegans* model for disease, we here identified MOAG-2/LIR-3 as regulator of RNA Pol III transcription that — in presence of polyglutamine expansion proteins — turns into a positive regulator of polyglutamine aggregation.

The role of MOAG-2/LIR-3 as a transcription factor has been previously suggested by others [23, 32, 33]. Its initial discovery as a putative transcription factor was due to its structural similarity to the C2H2 zinc fingers of LIN-26, a fate regulator responsible for the differentiation of non-neuronal ectodermal cells and somatic gonad epithelium [23, 42, 43]. By combining ChIP seq with RNA sequencing, we provide the first experimental evidence of MOAG-2/LIR-3 being a regulator of transcription.

In *C. elegans*, several transcription factors have been shown to bind to more than 10% to the promoters of small ncRNAs, but only one (GEI-11) is predicted to regulate ncRNAs [41]. We showed that MOAG-2/LIR-3 is required for the transcription of snRNA, snoRNA and tRNA genes and — unlike the majority of the other transcription factors — MOAG-2/LIR-3 bound to the same target genes as the RNA Pol III complex.

MOAG-2/LIR-3 expression in *C. elegans* has been described to localize to the nucleus of body wall muscle cells, the vulval muscles, the spermatheca, the head and tail ganglia and the ventral nerve cord, from embryogenesis throughout adulthood (www.transgeneome.mpi-cbg.de, March 2014 and [32, 33]). A microarray performed in two distinct mechanosensory neurons — the touch receptor neurons and the FLP neurons — revealed that *moag-2/lir-3* was upregulated in the FLP sensory neurons, suggesting that *moag-2/lir-3* is required for FLP differentiation [44]. Because the expression of MOAG-2/LIR-3 is limited to a subset of cell types, this suggests that MOAG-2/LIR-3 could be a tissue-specific regulator of transcription rather than a core component of the RNA Pol III machinery. Such cell type specific regulators of RNA Pol III have been described for human cells and proposed to accommodate tissue-specific needs for small non-coding RNAs ([45, 46], also reviewed in [39, 47]). Whether MOAG-3/LIR-3 has a similar role remains to be established.

In *C. elegans*, the nuclear pore protein NPP-13 has been described to associate with the RNA Pol III complex, which regulates the efficient processing of snoRNA and tRNA transcripts [36]. Knockdown of NPP-13 results in abnormally long snoRNA and tRNA transcripts that cannot be processed into their mature form [36]. We did not find unprocessed transcripts neither for snoRNA nor tRNAs in *moag-2/lir-3* mutant animals (data not shown), which excludes the possibility that MOAG-2/LIR-3 is required for RNA Pol III transcript processing. Moreover, knockdown of NPP-13 did not alter aggregation, which suggests that mutations in *moag-2/lir-3* do not alter aggregation by interfering with the nuclear pore complex.

In this study, expression of an aggregation-prone protein downregulated the levels of snRNAs, snoRNAs and tRNAs, demonstrating that aggregation-prone proteins also affect the non-coding genome. One explanation for this downregulation is that the aggregation-prone proteins sequestered or changed the localization of MOAG-2/LIR-3, and perhaps also components of the RNA Pol III complex. Several intrinsically disordered proteins, which include proteins involved in transcriptional regulation, are known to be sequestered by aggregation-prone proteins into aggregates [3, 48-51]. One example

is Sp1, which can no longer bind to its DNA targets due to sequestration by mutant *huntingtin* [52, 53]. Because MOAG-2/LIR-3 is a small (284 aa), predominantly disordered protein (75.5%) (<http://bip.weizmann.ac.il/fldbin/findex>, June 2015), its sequestration could explain the lower levels of small ncRNAs observed in presence of aggregation-prone proteins. Nevertheless, the fact that aggregation-prone proteins affect small ncRNA expression poses the interesting question whether this is correlated their toxicity. It will also be interesting to find whether the cellular downregulation of small ncRNAs is a consequence specific for polyglutamine expansion proteins or shared by other aggregation-prone proteins such as alpha-synuclein and amyloid-beta.

Individual knockdown of the components of the RNA Pol III complex did not affect aggregation in Q40 animals, indicating that RNA Pol III transcription is not related to the effect of *moag-2/lir-3* mutations on protein aggregation. Impaired processing of precursor tRNA into its mature form or mutations in the tRNA body sequence have been implicated in neurodegeneration [54-59]. In *C. elegans*, mutations in the tRNA body sequence can result in protein misfolding, protein aggregation and neurological defects [60, 61]. We found, however, that knockdown of tRNA processing enzymes as well as factors known to cooperate with those enzymes (*rtcb-1*, *clpf-1* and *gtbp2* homolog) did not alter the aggregation phenotype in Q40 animals, also indicating that impaired tRNA processing is not the cause for protein aggregation in our model. The glutaminyl tRNA that specifically recognizes the CAG codon is required in high demand during translation of the expanded CAG tract in *huntingtin*, eventually resulting in its depletion [62]. According to our data, there is no evidence for a preferential enrichment or depletion of glutaminyl pre-tRNAs in Q40 animals (data not shown). In addition, while we cannot account for mature tRNAs present in the body wall muscle cells, which would require detection in a tissue-specific manner, the polyglutamine expression levels are not decreased in the *moag-2/lir-3* mutant animals, reducing the possibility that the availability of glutaminyl tRNAs plays a role.

Several transcription factors have been implicated in protein aggregation, but none with known function related to RNA Pol III transcription [53, 63]. Our results indicate that MOAG-2/LIR-3 regulated protein aggregation independently of its

role in small ncRNA transcription and processing, as shown by RNAi knockdown experiments. One possibility is that MOAG-2/LIR-3 drives aggregation directly. Precedents for aggregation-promoting factors include SERF1A, which can drive aggregation of a variety of disease proteins via a transient direct interaction with early aggregation intermediates and SH3GL3, which was shown to bind physically to the *huntingtin* protein or its mutated form through the SH3 domain [11, 12, 14, 15].

Components of the RNA Pol II complex have been reported to be part of protein aggregates. Such an example comes from TBP, which was localized to nuclear inclusions of SCA-1, -2, -3 and neurofibrillary tangles of Alzheimer's disease [64-66]. A regulatory role for TBP in protein aggregation, as we observe here for MOAG-2/LIR-3, however, has not been reported.

In summary, this work reveals that protein aggregation can affect the non-coding genome by altering the expression of ncRNA genes available in the cell. Secondly, this work opens yet another perspective on how aggregation-prone disease proteins can impair cellular homeostasis by inhibiting transcription of non-coding RNAs and converting normal cellular proteins into aggregation-promoting factors at the expense of their function.

Methods

Strains and genetics

Standard methods were used for culturing *C. elegans* at 20°C [67]. The LIR-3::GFP strain (OP312) was generated by biolistic transformation to produce an integrated, low-copy transgene of the WRM0637aB05 fosmid, recombineered with GFP:3XFLAG in frame at the carboxy terminus of the *lir-3* locus [34]. To synchronize animals, eggs were collected from gravid hermaphrodites by hypochlorite bleaching and hatched overnight in M9 buffer. The desired amount of L1 animals were subsequently cultured on nematode growth medium agar plates (#633185, Greiner Bio-One) seeded with OP50 bacteria. The following strains were used or generated: wild type N2 (Bristol), AM141 *rmls133[P(unc-54)Q40::YFP]X* [68], OP312 *wgl312[P(lir-3)::TY1::EGFP::3xFLAG + unc-119(+)]* [35, 69], OW1002 *lir-3(tm0813)II*, OW1003 *rmls133[P(unc-54)Q40::YFP]X;lir-3(tm813)II*, OW1004 *rmls133[P(unc-54)Q40::YFP]X*, OW1019

moag-2 (pk2183), OW1020 *rmls133[P(unc-54)Q40::YFP]X;moag-2(pk2183)II*, OW1021 *rmls133[P(unc-54)Q40::YFP]X*, OW1086 *rmls133[P(unc-54)::Q40::YFP]X*, OW1087 *rmls133[P(unc-54)::Q40::YFP]X*; *zgEx221[P(myo-3)::CFP + P(lir-3)::lir-3]; lir-3(tm813)II*, OW1090 *rmls133[P(unc-54)::Q40::YFP]X*, OW1091 *rmls133[P(unc-54)::Q40::YFP]X*; *zgEx226[P(myo-3)::CFP + P(lir-3)::lir-3]; lir-3(tm813)II*, OW1100 *rmls133[P(unc-54)::Q40::YFP]X*; *wgls312[P(lir-3)::TY1::EGFP::3xFLAG + unc-119(+)]*.

EMS Mutagenesis and Mapping

Mutagenesis was performed using standard *C. elegans* ethyl methanesulfonate (EMS) methodology [70]. Eight thousand mutagenized genomes were screened for suppressors of aggregation. *moag-2(pk2183)* was identified by single-nucleotide polymorphism mapping to a region between base 9,400,743 and 11,827,697 on linkage group II [71]. Next generation sequencing was performed in that region to identify candidate genes for *moag-2*. CLC Bio (Qiagen, <http://www.clcbio.com>) and MAQGene softwares were utilized for mapping the mutation in F37H8.1 [72].

Creation of transgenic strains

For the rescue experiment, a genomic construct of *lir-3* spanning 1500bp upstream to 330bp downstream of F37H8.1 was amplified from N2 genomic DNA by nested PCR using primers F1: CGCTCACAGTCAACGTCG; R1: CCATGCGATTTGACACATTTTCG; F2: CGGCATTGCTCTTGTCGTGC and R2: GCATCTCATGAAACCAGACGC. The resulting PCR fragment was cloned into the pGEM-T Easy Vector (Promega) and sequenced. Transgenic lines were made by injecting ~20 ng/ul of construct along with ~10 ng/ul of pPD136.61 [*P(unc-54)::CFP*] in N2 animals.

RNAi experiments

RNAi experiments were performed on NGM agar plates containing 1 mM IPTG and 50 mg/ml ampicillin that were seeded with RNAi bacteria induced with IPTG to produce dsRNA. Animals were synchronized by hypochlorite bleaching and L1 animals were grown on RNAi plates (#628103, Greiner Bio-One) and used for the experiments at L4 stage, unless stated otherwise.

Quantification of aggregates

The number of aggregates present in whole animals was counted using a fluorescence dissection microscope (Leica Microsystems, Wetzlar, Germany). A minimum of 20 animals was counted in three biological replicates, unless stated otherwise.

Filter retardation assay

The protocol was adapted from Wanker et al [27, 28]. Briefly, crude worm lysates from synchronized L4 animals were resuspended in FTA Sample Buffer (10 mM Tris-Cl pH 8.0; 150 mM NaCl; 2% SDS) and disrupted using a bead-beater (FastPrep 24, MP Biomedicals) for 7 cycles of 20 seconds bead beating/5 minutes rest per cycle. Supernatants were transferred to new 1.5 ml tubes and protein concentration was determined using the BCA Protein Assay Kit (#23227; Thermo Scientific Pierce). To detect SDS-insoluble aggregates, 100 ug of total protein was mixed with 1 M DTT and FTA Sample buffer (final concentration 40 ug/100 ul) and heated for 98°C for 5 minutes. Samples were filtered through a 0.22 micron cellulose acetate membrane using a Bio-Dot microfiltration apparatus (Biorad) and 100 ug of total protein was used followed by five-fold serial dilutions. Proteins were blocked for 30 minutes with 5% milk in TBS-T. Membranes were incubated with GFP (#632381, Clontech) or α -tubulin (#T6074-200UL, Sigma) primary antibodies at a 1:5000 dilution overnight at 4°C. Incubation with secondary antibody anti-mouse was applied at a 1:10.000 dilution for 1 hour at room temperature. Antibody binding was visualized with an ECL kit (#RPN2232, Amersham).

Chromatin immunoprecipitation sequencing

ChIP assays were conducted as previously described [35, 41]. Worm staging was achieved by bleaching and L1 starvation. Arrested L1 animals were plated on peptone-enriched NGM plates seeded with OP50 bacteria and grown for 48 hours for L4 collection at 20°C. Samples were crosslinked with 2% formaldehyde for 30 minutes at room temperature and then quenched with 1 M Tris pH 7.5. The pelleted worms were subsequently flash frozen in liquid nitrogen and stored at -80°C. Samples were sonicated using a microtip to obtain mostly 200 to 800 bp DNA fragments. For each sample, 2.2 or 4.4 mg of cell

extract was immunoprecipitated using a goat anti-GFP, GoatV (gift from Kevin White). The enriched DNA fragments and input control (genomic DNA from the same sample) for two biological replicates were used for library preparation and sequencing as previously described [73]. Briefly, samples were librated and multiplexed using the Ovation Ultralow DR Multiplex Systems 1-8 and 9-16 (NuGEN Technologies Inc., San Carlos, CA) following the manufacturer's protocol except Qiagen MinElute PCR purification kits were used to isolate the DNA. Library size selection in the 200-800 bp range was achieved using the SPRIselect reagent kit (Beckman Coulter, Inc., Brea, CA) and sequencing was performed on the Illumina HiSeq 2000 platform. To search for MOAG-2/LIR-3 specific binding sites, only binding sites consistent across both replicates and within a range of -400 bp to +100 bp distance from TSS were considered. In addition, we excluded highly occupied target regions to avoid false positives in our analysis [74]. The ChIP seq data used was obtained from modENCODE DCC (<http://www.modencode.org>; September 2014). All ChIP-seq data have been deposited in Gene Expression Omnibus (GEO), under accession number (in preparation).

RNA sequencing

Worms were grown to L4 stage and total RNA was extracted using TRIzol Reagent (Life Technologies) according to the manufacturer's description. For polyA RNA sequencing, the TruSeq Sample Preparation V2 Kit was used (Illumina). For non-coding RNA sequencing (mRNA, snoRNA, tRNA and large ncRNA), a protocol has been described previously [36]. Briefly, total RNA was treated with DNaseI and depleted from rRNA with the Ribominus Eukaryote Kit (#A10837-08, Invitrogen). Fragmentation of RNA was performed using a Fragmentation Buffer (#AM8740, Ambion) and cDNA was generated using the Superscript II Kit (Invitrogen). cDNA libraries were subjected to high-throughput single-end sequencing (50 bp) in an Illumina HiSeq 2500 instrument. The RNA sequencing dataset has been submitted to GEO, under accession number (in preparation). RNA-sequencing data was mapped to WS220 genome reference using TopHat 2.0.9 program [75] and gene annotation from Ensembl release 66. Per gene expression data was normalized as fragments per million mapped (FPM). Data visualization and statistical tests were conducted using R scripts.

Bioinformatic Analysis

Conserved domains were identified using SMART (Simple Modular Architecture Research Tool) [76, 77]. Prediction of nuclear signal localization was done with NLS Mapper [78]. Prediction of protein folding was done with FoldIndex [79]. The algorithm used for motif discovery was The MEME Suite [80]. Orthologs were identified using protein BLAST search and aligned with T-Coffee multiple sequence alignment tool [81]. Amino acid predictions were performed using ORF Finder (<http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi>).

Acknowledgements

We thank the Caenorhabditis Genetics Centre (funded by the National Institutes of Health National Centre for Research Resources) and the Mitani laboratory for the *C. elegans* strains (funded by the Japan National BioResource Project). We thank Pieter van der Vlies and Jelke Bergsma for practical assistance with the whole-genome sequencing. We thank Pieter Neerincx for help with installing MAQGene. We thank Inge Kaziemer, Nancy Halsema, Marianna Bevova and Diana Spierings for assisting with RNA sequencing. We thank Kohta Ikegami for the *npp3* RNAi vector. We thank the modERN project (NIH U41HG007355) for sharing the MOAG-2/LIR-3 ChIP seq data prior to publication. Alumni chapter Gooische Groningers facilitated by the Ubbo Emmius Fund. This project was funded by a European Research Council (ERC) starting grant (to E.A.A.N.), the Fundação para a Ciência e Tecnologia fellowship (SFRH/BPD/51009/2010) (to O.S.), a Groningen University Institute for Drug Exploration (GUIDE) fellowship (O.S.).

Author Contributions

O.S. and E.A.A.N. designed the research; O.S., M.K., R.S., A.M.C., C.M., A. Z., H.H.W. and W.H. performed the research; O.S., T.v.d.J., R.M.P., V.G. analyzed the data; and O.S. and E.A.A.N. wrote the manuscript.

References

1. Soto C (2003) Unfolding the role of protein misfolding in neurodegenerative diseases. *Nat Rev Neurosci* 4:49–60. doi: 10.1038/nrn1007
2. Eisenberg D, Jucker M (2012) The amyloid state of proteins in human diseases. *Cell* 148:1188–1203. doi: 10.1016/j.cell.2012.02.022
3. Olzscha H, Schermann SM, Woerner AC, et al. (2011) Amyloid-like Aggregates Sequester Numerous Metastable Proteins with Essential Cellular Functions. *Cell* 144:67–78. doi: 10.1016/j.cell.2010.11.050
4. Chiti F, Dobson CM (2009) Amyloid formation by globular proteins under native conditions. *Nat Chem Biol* 5:15–22. doi: 10.1038/nchembio.131
5. Escusa-Toret S, Vonk WIM, Frydman J (2013) Spatial sequestration of misfolded proteins by a dynamic chaperone pathway enhances cellular fitness during stress. *Nat Cell Biol* 15:1231–1243. doi: 10.1038/ncb2838
6. Kaganovich D, Kopito R, Frydman J (2008) Misfolded proteins partition between two distinct quality control compartments. *Nature* 454:1088–1095. doi: 10.1038/nature07195
7. Miller J, Arrasate M, Brooks E, et al. (2011) Identifying polyglutamine protein species in situ that best predict neurodegeneration. *Nat Chem Biol* 7:925–934. doi: 10.1038/nchembio.694
8. Bolognesi B, Kumita JR, Barros TP, et al. (2010) ANS Binding Reveals Common Features of Cytotoxic Amyloid Species. *5:735–740*. doi: 10.1021/cb1001203
9. Arrasate M, Mitra S, Schweitzer ES, et al. (2004) Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* 431:805–810. doi: 10.1038/nature02998
10. Kaye R, Head E, Thompson JL, et al. (2003) Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* 300:486–489. doi: 10.1126/science.1079469
11. Falsone SF, Meyer NH, Schrank E, et al. (2012) SERF protein is a direct modifier of amyloid fiber assembly. *Cell Rep* 2:358–371. doi: 10.1016/j.celrep.2012.06.012
12. van Ham TJ, Holmberg MA, van der Goot AT, et al. (2010) Identification of MOAG-4/SERF as a regulator of age-related proteotoxicity. *Cell* 142:601–612. doi: 10.1016/j.cell.2010.07.020
13. Schaffar G, Breuer P, Boteva R, et al. (2004) Cellular toxicity of polyglutamine expansion proteins: mechanism of transcription factor deactivation. *Molecular Cell* 15:95–105. doi: 10.1016/j.molcel.2004.06.029
14. Sittler A, Wälter S, Wedemeyer N, et al. (1998) SH3GL3 associates with the Huntingtin exon 1 protein and promotes the formation of polyglutamine-containing protein aggregates. *Molecular Cell* 2:427–436.
15. Davranche A, Aviat H, Zeder-Lutz G, et al. (2011) Huntingtin affinity for partners is not changed by polyglutamine length: aggregation itself triggers aberrant interactions. *Human Molecular Genetics* 20:2795–2806. doi: 10.1093/hmg/ddr178
16. Garcia SM, Casanueva MO, Silva MC, et al. (2007) Neuronal signaling modulates protein homeostasis in *Caenorhabditis elegans* post-synaptic muscle cells. *Genes & Development* 21:3006–3016. doi: 10.1101/gad.1575307
17. van der Goot AT, Zhu W, Vazquez-Manrique RP, et al. (2012) Delaying aging and the aging-associated decline in protein homeostasis by inhibition of tryptophan degradation. *Proc Natl Acad Sci USA* 109:14912–14917. doi: 10.1073/pnas.1203083109
18. Lejeune F-X, Mesrob L, Parmentier F, et al. (2012) Large-scale functional RNAi screen in *C. elegans* identifies genes that regulate the dysfunction of mutant polyglutamine neurons. *BMC Genomics* 13:91. doi: 10.1186/1471-2164-13-91
19. Silva MC, Fox S, Beam M, et al. (2011) A genetic screening strategy identifies novel regulators of the proteostasis network. *PLoS Genet* 7:e1002438. doi: 10.1371/journal.pgen.1002438
20. Hamamichi S, Rivas RN, Knight AL, et al. (2008) Hypothesis-based RNAi screening identifies neuroprotective genes in a Parkinson's disease model. *Proc Natl Acad Sci USA* 105:728–733. doi: 10.1073/pnas.0711018105
21. Kuwahara T, Koyama A, Koyama S, et al. (2008) A systematic RNAi screen reveals involvement of endocytic pathway in neuronal dysfunction in alpha-synuclein transgenic *C. elegans*. *Human Molecular Genetics* 17:2997–3009. doi: 10.1093/hmg/ddn198
22. Nollen EAA, Garcia SM, van Haaften G, et al. (2004) Genome-wide RNA interference screen identifies previously undescribed regulators of polyglutamine aggregation. *Proceedings of the*

- National Academy of Sciences 101:6403–6408. doi: 10.1073/pnas.0307697101
23. Dufourcq P, Chanal P, Vicaire S, et al. (1999) *lir-2*, *lir-1* and *lin-26* encode a new class of zinc-finger proteins and are organized in two overlapping operons both in *Caenorhabditis elegans* and in *Caenorhabditis briggsae*. *Genetics* 152:221–235.
 24. Yanamandra K, Jiang H, Mahan TE, et al. (2015) Anti-tau antibody reduces insoluble tau and decreases brain atrophy. *Ann Clin Transl Neurol* 2:278–288. doi: 10.1002/acn3.176
 25. Shankar GM, Li S, Mehta TH, et al. (2008) Amyloid-beta protein dimers isolated directly from Alzheimer's brains impair synaptic plasticity and memory. *Nat Med* 14:837–842. doi: 10.1038/nm1782
 26. Lee VM, Wang J, Trojanowski JQ (1999) Purification of paired helical filament tau and normal tau from human brain tissue. *Meth Enzymol* 309:81–89.
 27. Wanker EE, Scherzinger E, Heiser V, et al. (1999) Membrane filter assay for detection of amyloid-like polyglutamine-containing protein aggregates. *Meth Enzymol* 309:375–386.
 28. Scherzinger E, Lurz R, Turmaine M, et al. (1997) Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. *Cell* 90:549–558.
 29. Brown RS (2005) Zinc finger proteins: getting a grip on RNA. *Curr Opin Struct Biol* 15:94–98. doi: 10.1016/j.sbi.2005.01.006
 30. Hall TMT (2005) Multiple modes of RNA recognition by zinc finger proteins. *Curr Opin Struct Biol* 15:367–373. doi: 10.1016/j.sbi.2005.04.004
 31. Krishna SS, Majumdar I, Grishin NV (2003) Structural classification of zinc fingers: survey and summary. *Nucleic Acids Research* 31:532–550. doi: 10.1093/nar/gkg161
 32. Reece-Hoyes JS, Shingles J, Dupuy D, et al. (2007) Insight into transcription factor gene duplication from *Caenorhabditis elegans* Promoterome-driven expression patterns. *BMC Genomics* 8:27. doi: 10.1186/1471-2164-8-27
 33. Reece-Hoyes JS, Deplancke B, Shingles J, et al. (2005) A compendium of *Caenorhabditis elegans* regulatory transcription factors: a resource for mapping transcription regulatory networks. *Genome Biol* 6:R110. doi: 10.1186/gb-2005-6-13-r110
 34. Sarov M, Murray JI, Schanze K, et al. (2012) A genome-scale resource for in vivo tag-based protein function exploration in *C. elegans*. *Cell* 150:855–866. doi: 10.1016/j.cell.2012.08.001
 35. Zhong M, Niu W, Lu ZJ, et al. (2010) Genome-Wide Identification of Binding Sites Defines Distinct Functions for *Caenorhabditis elegans* PHA-4/FOXA in Development and Environmental Response. *PLoS Genet* 6:e1000848. doi: 10.1371/journal.pgen.1000848
 36. Ikegami K, Lieb JD (2013) Integral Nuclear Pore Proteins Bind to Pol III-Transcribed Genes and Are Required for Pol III Transcript Processing in *C. elegans*. *Molecular Cell* 51:840–849. doi: 10.1016/j.molcel.2013.08.001
 37. Schramm L, Hernandez N (2002) Recruitment of RNA polymerase III to its target promoters. *Genes & Development* 16:2593–2620. doi: 10.1101/gad.1018902
 38. Bratkovič T, Rogelj B (2014) The many faces of small nucleolar RNAs. *Biochim Biophys Acta* 1839:438–443. doi: 10.1016/j.bbaggm.2014.04.009
 39. White RJ (2011) Transcription by RNA polymerase III: more complex than we thought. *Nature Reviews Genetics* 12:459–463. doi: 10.1038/nrg3001
 40. Guthrie C, Patterson B (1988) Spliceosomal snRNAs. *Annu Rev Genet* 22:387–419. doi: 10.1146/annurev.ge.22.120188.002131
 41. Niu W, Lu ZJ, Zhong M, et al. (2011) Diverse transcription factor binding features revealed by genome-wide ChIP-seq in *C. elegans*. *Genome Research* 21:245–254. doi: 10.1101/gr.114587.110
 42. Boer den BG, Sookhareea S, Dufourcq P, Labouesse M (1998) A tissue-specific knockout strategy reveals that *lin-26* is required for the formation of the somatic gonad epithelium in *Caenorhabditis elegans*. *Development* 125:3213–3224.
 43. Labouesse M, Hartwig E, Horvitz HR (1996) The *Caenorhabditis elegans* LIN-26 protein is required to specify and/or maintain all non-neuronal ectodermal cell fates. *Development* 122:2579–2588.
 44. Topalidou I, Chalfie M (2011) Shared gene expression in distinct neurons expressing common selector genes. *Proc Natl Acad Sci USA* 108:19258–19263. doi: 10.1073/pnas.1111684108
 45. Alla RK, Cairns BR (2014) RNA polymerase III transcriptomes in human embryonic stem cells and induced pluripotent stem cells, and relationships with pluripotency transcription

- factors. PLoS ONE 9:e85648. doi: 10.1371/journal.pone.0085648
46. Oler AJ, Alla RK, Roberts DN, et al. (2010) Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. *Nature Publishing Group* 17:620–628. doi: 10.1038/nsmb.1801
47. Marshall L, White RJ (2008) Non-coding RNA production by RNA polymerase III is implicated in cancer. *Nat Rev Cancer* 8:911–914. doi: 10.1038/nrc2539
48. Walther DM, Kasturi P, Zheng M, et al. (2015) Widespread Proteome Remodeling and Aggregation in Aging *C. elegans*. *Cell* 161:919–932. doi: 10.1016/j.cell.2015.03.032
49. Liu J, Perumal NB, Oldfield CJ, et al. (2006) Intrinsic disorder in transcription factors. *Biochemistry* 45:6873–6888. doi: 10.1021/bi0602718
50. Minezaki Y, Homma K, Kinjo AR, Nishikawa K (2006) Human transcription factors contain a high fraction of intrinsically disordered regions essential for transcriptional regulation. *Journal of Molecular Biology* 359:1137–1149. doi: 10.1016/j.jmb.2006.04.016
51. Iakouchcheva LM, Brown CJ, Lawson JD, et al. (2002) Intrinsic Disorder in Cell-signaling and Cancer-associated Proteins. *Journal of Molecular Biology* 323:573–584. doi: 10.1016/S0022-2836(02)00969-5
52. Dunah AW, Jeong H, Griffin A, et al. (2002) Sp1 and TAFII130 transcriptional activity disrupted in early Huntington's disease. *Science* 296:2238–2243. doi: 10.1126/science.1072613
53. Li S-H, Cheng AL, Zhou H, et al. (2002) Interaction of Huntington disease protein with transcriptional activator Sp1. *Molecular and Cellular Biology* 22:1277–1287. doi: 10.1128/MCB.22.5.1277-1287.2002
54. Ishimura R, Nagy G, Dotu I, et al. (2014) RNA function. Ribosome stalling induced by mutation of a CNS-specific tRNA causes neurodegeneration. *Science* 345:455–459. doi: 10.1126/science.1249749
55. Schaffer AE, Eggens VRC, Caglayan AO, et al. (2014) CLP1 Founder Mutation Links tRNA Splicing and Maturation to Cerebellar Development and Neurodegeneration. *Cell* 157:651–663. doi: 10.1016/j.cell.2014.03.049
56. Hanada T, Weitzer S, Mair B, et al. (2013) CLP1 links tRNA metabolism to progressive motor-neuron loss. *Nature* 1–7. doi: 10.1038/nature11923
57. Karaca E, Weitzer S, Pehlivan D, et al. (2014) Human CLP1 Mutations Alter tRNA Biogenesis, Affecting Both Peripheral and Central Nervous System Function. *Cell* 157:636–650. doi: 10.1016/j.cell.2014.02.058
58. Jordanova A, Irobi J, Thomas FP, et al. (2006) Disrupted function and axonal distribution of mutant tyrosyl-tRNA synthetase in dominant intermediate Charcot-Marie-Tooth neuropathy. *Nat Genet* 38:197–202. doi: 10.1038/ng1727
59. Lee JW, Beebe K, Nangle LA, et al. (2006) Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration. *Nature* 443:50–55. doi: 10.1038/nature05096
60. Nedialkova DD, Leidel SA (2015) Optimization of Codon Translation Rates via tRNA Modifications Maintains Proteome Integrity. *Cell*. doi: 10.1016/j.cell.2015.05.022
61. Chen C, Tuck S, Byström AS (2009) Defects in tRNA Modification Associated with Neurological and Developmental Dysfunctions in *Caenorhabditis elegans* Elongator Mutants. *PLoS Genet* 5:e1000561–15. doi: 10.1371/journal.pgen.1000561
62. Girstmair H, Saffert P, Rode S, et al. (2013) Depletion of cognate charged transfer RNA causes translational frameshifting within the expanded CAG stretch in huntingtin. *Cell Rep* 3:148–159. doi: 10.1016/j.celrep.2012.12.019
63. Hsu A-L, Murphy CT, Kenyon C (2003) Regulation of aging and age-related disease by DAF-16 and heat-shock factor. *Science* 300:1142–1145. doi: 10.1126/science.1083701
64. Uchihara T, Fujigasaki H, Koyano S, et al. (2001) Non-expanded polyglutamine proteins in intranuclear inclusions of hereditary ataxias—triple-labeling immunofluorescence study. *Acta Neuropathol* 102:149–152. doi: 10.1007/s004010100364
65. Reid SJ, van Roon-Mom WMC, Wood PC, et al. (2004) TBP, a polyglutamine tract containing protein, accumulates in Alzheimer's disease. *Molecular Brain Research* 125:120–128. doi: 10.1016/j.molbrainres.2004.03.018
66. Perez MK, Paulson HL, Pendse SJ, et al. (1998) Recruitment and the role of nuclear localization in polyglutamine-mediated aggregation. *J Cell Biol* 143:1457–1470.
67. Brenner S (1974) The genetics of *Caenorhabditis elegans*. *Genetics* 77:71–94.
68. Morley JF, Brignull HR, Weyers JJ, Morimoto RI (2002) The threshold for polyglutamine-expansion protein aggregation and cellular

- toxicity is dynamic and influenced by aging in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences* 99:10417–10422. doi: 10.1073/pnas.152161099
69. Sarov M, Schneider S, Pozniakovski A, et al. (2006) A recombineering pipeline for functional genomics applied to *Caenorhabditis elegans*. *Nature Methods* 3:839–844. doi: 10.1038/nmeth933
 70. Jorgensen EM, Mango SE (2002) The art and design of genetic screens: *caenorhabditis elegans*. *Nature Reviews Genetics* 3:356–369. doi: 10.1038/nrg794
 71. Wicks SR, Yeh RT, Gish WR, et al. (2001) Rapid gene mapping in *Caenorhabditis elegans* using a high density polymorphism map. *Nat Genet* 28:160–164. doi: 10.1038/88878
 72. Bigelow H, Doitsidou M, Sarin S, Hobert O (2009) MAQGene: software to facilitate *C. elegans* mutant genome sequence analysis. *Nature Methods* 6:549–549. doi: 10.1038/nmeth.f.260
 73. Kasper DM, Wang G, Gardner KE, et al. (2014) The *C. elegans* SNAPc component SNPC-4 coats piRNA domains and is globally required for piRNA abundance. *Developmental Cell* 31:145–158. doi: 10.1016/j.devcel.2014.09.015
 74. Gerstein MB, Lu ZJ, Van Nostrand EL, et al. (2010) Integrative Analysis of the *Caenorhabditis elegans* Genome by the modENCODE Project. *Science* 330:1775–1787. doi: 10.1126/science.1196914
 75. Kim D, Pertea G, Trapnell C, et al. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14:R36. doi: 10.1186/gb-2013-14-4-r36
 76. Schultz J, Milpetz F, Bork P, Ponting C (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proceedings of the National Academy of Sciences* 95:5857–5864. doi: 10.1073/pnas.95.11.5857
 77. Letunic I, Doerks T, Bork P (2012) SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Research* 40:D302–5. doi: 10.1093/nar/gkr931
 78. Kosugi S, Hasebe M, Tomita M, Yanagawa H (2009) Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proceedings of the National Academy of Sciences* 106:10171–10176. doi: 10.1073/pnas.0900604106
 79. Brundin P, Melki R, Kopito R (2010) Prion-like transmission of protein aggregates in neurodegenerative diseases. *Nature Publishing Group* 11:301–307. doi: 10.1038/nrm2873
 80. Bailey TL, Bodén M, Buske FA, et al. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research* 37:W202–8. doi: 10.1093/nar/gkp335
 81. Notredame C, Higgins DG, Heringa J (2000) T-coffee: a novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology* 302:205–217. doi: 10.1006/jmbi.2000.4042

Supplemental figures

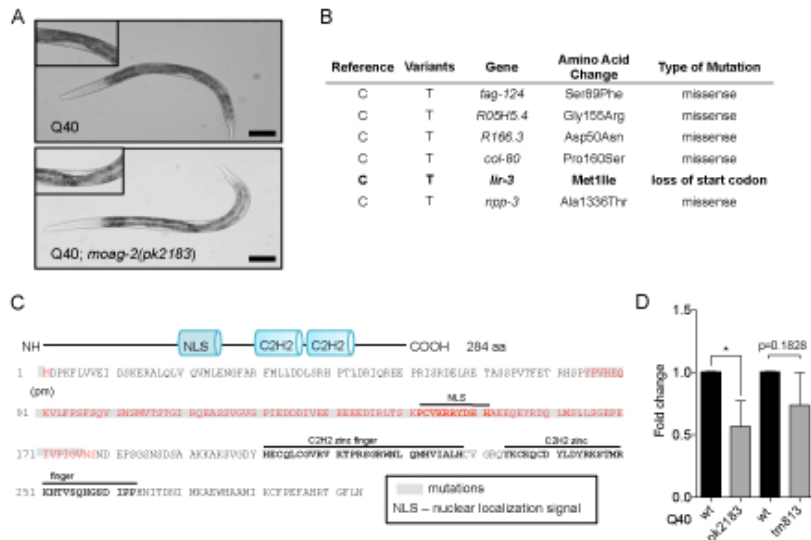


Figure S1. (A) Representative bright field images of L4-staged Q40 and Q40;*lir-3(pk2183)* animals. Insets show a magnification of the vulva. Scale bar, 75 μ m. **(B)** List of candidate genes for *moag-2* obtained by whole genome sequencing. **(C)** Amino acid sequence of LIR-3 with the predicted nuclear localization signal (NLS, residues 132-141) and C2H2 zinc finger domains (residues 191-214 and 224-247). Grey shadow indicates location of the point mutation and the deleted residues in the *lir-3(tm813)* mutant. **(D)** Quantification of filter retardation assays. Immunoblots were quantified by densitometry using ImageJ. Presented is the ratio (fold change) between protein levels of Q40;*moag-2/lir-3* mutants relative to their corresponding wild types (corrected to α -tubulin as a loading control). Filter retardation assay was performed at the L4 stage and represented is the average of three biological replicates. Data are represented as mean \pm SEM and significance was calculated using one-tailed unpaired t-test. * $p < 0.05$

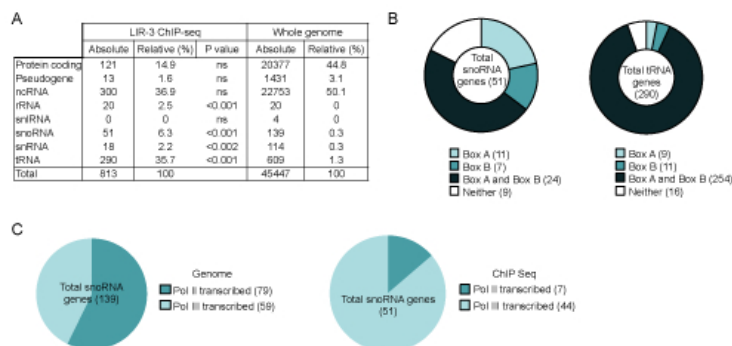


Figure S2. (A) Number of MOAG-2/LIR-3 binding sites per gene biotype and total number of known *C. elegans* genes. p value was calculated by permutation test. **(B)** Number of MOAG-2/LIR-3 binding sites containing Box A and Box B in snoRNA and tRNA genes. **(C)** Number of snoRNA genes encoded by RNA Pol II or III genome-wide and detected in this study.

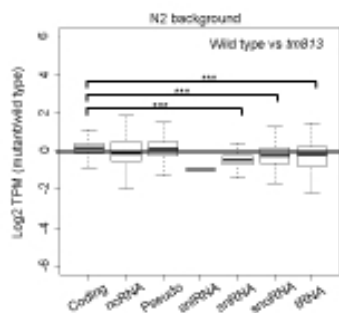


Figure S3. Boxplot showing the relative expression of different gene biotypes in *lir-3(tm813)* relative to the wild type N2 background. Coding: protein coding genes; ncRNA: non-coding RNA; Pseudo: pseudogenes; snRNA: small nuclear RNA; snlRNA: snRNA-like RNA, snoRNA: small nucleolar RNA; tRNA: transfer RNA. *** $p < 0.001$

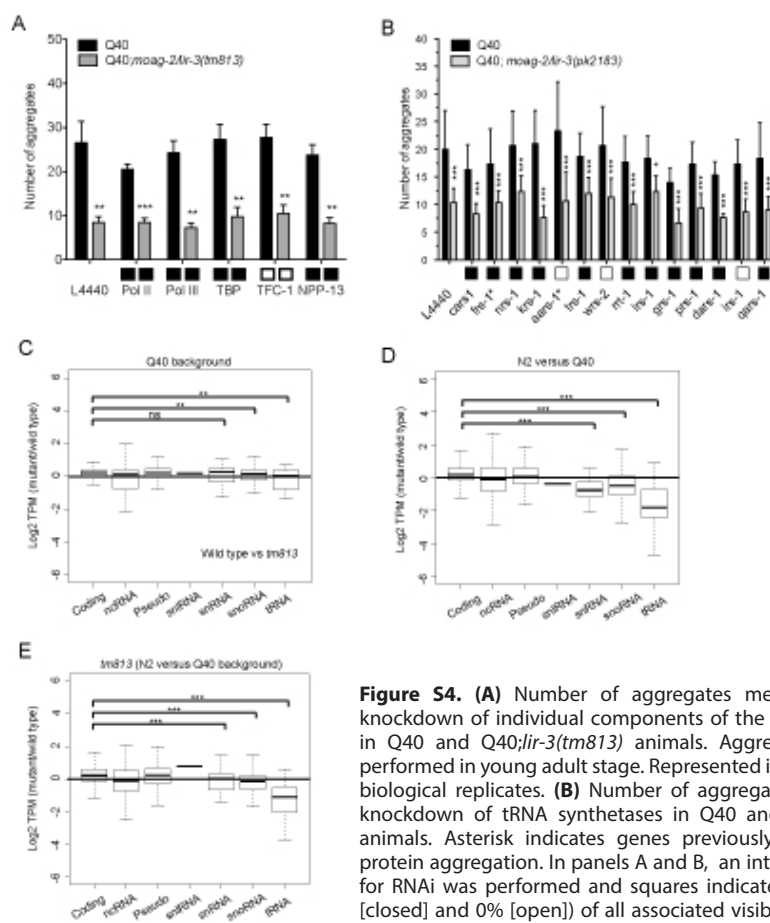


Figure S4. (A) Number of aggregates measured upon RNAi knockdown of individual components of the RNA Pol III complex in Q40 and Q40;*lir-3(tm813)* animals. Aggregate counting was performed in young adult stage. Represented is the average of four biological replicates. (B) Number of aggregates measured upon knockdown of tRNA synthetases in Q40 and Q40;*lir-3(pk2183)* animals. Asterisk indicates genes previously known to reduce protein aggregation. In panels A and B, an internal quality control for RNAi was performed and squares indicate penetrance (100% [closed] and 0% [open]) of all associated visible RNAi phenotypes other than aggregation. Data are represented as mean \pm SEM and significance was calculated using two-tailed unpaired t-test. (C-E) Boxplot showing the relative expression of different gene biotypes in (C) *lir-3(tm813)* relative to its wild type in the Q40 background; in (D) in the N2 and Q40 wild type animals (Q40 control outcrossed from *lir-3(tm813)*); and in (E) *lir-3(tm813)* in the N2 and Q40 genetic backgrounds. Coding: protein coding genes; ncRNA: non-coding RNA; Pseudo: pseudogenes; snRNA: small nuclear RNA; snlRNA: snRNA-like RNA, snoRNA: small nucleolar RNA; tRNA: transfer RNA. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ns is not significant. In panels B to E, the average of three biological replicates is represented.

Table S1. Overview and statistical analysis of aggregate counting experiments.

Replicates (N)	Strain	Mean \pm SEM	versus	Mean \pm SEM	p value
Figure 1A					
N=3	Q40	33 \pm 0.85	Q40; <i>lir-3(pk2183)</i>	17 \pm 0.64	< 0.0001
Figure 1D					
N=3	Q40	49 \pm 1.0	Q40; <i>lir-3(tm813)</i>	31 \pm 0.72	< 0.0001
	Q40; <i>lir-3(het)</i>	47 \pm 1.0			< 0.0001
Figure 1E					
N=3	Q40	35 \pm 1.32	Q40 + <i>lir-3 overexpression</i>	43 \pm 1.26	p = 0.0001
	Q40; <i>lir-3(tm813)</i>	13 \pm 0.80	Q40; <i>lir-3(tm813)</i> + <i>lir-3 overexpression</i>	27 \pm 0.97	< 0.0001
Figure 4A					
N=3	Q40 on control RNAi	32 \pm 5.85	Q40; <i>lir-3(pk2183)</i> on control RNAi	14 \pm 4.99	p = 0.0393
	Q40 on <i>ama-1 (Pol II)</i>	20 \pm 1.50	Q40; <i>lir-3(pk2183)</i> on <i>ama-1 (Pol II)</i>	5 \pm 0.70	p = 0.0004
	Q40 on <i>rpc-1 (Pol III)</i>	34 \pm 6.65	Q40; <i>lir-3(pk2183)</i> on <i>rpc-1 (Pol III)</i>	15 \pm 5.85	p = 0.05
	Q40 on <i>tbp-1 (TBP)</i>	36 \pm 7.31	Q40; <i>lir-3(pk2183)</i> on <i>tbp-1 (TBP)</i>	16 \pm 6.20	p = 0.0479
	Q40 on <i>tftc-5 (TFC-1)</i>	39 \pm 8.19	Q40; <i>lir-3(pk2183)</i> on <i>tftc-5 (TFC-1)</i>	16 \pm 5.36	p = 0.0383
	Q40 on <i>npp-13</i>	34 \pm 6.89	Q40; <i>lir-3(pk2183)</i> on <i>npp-13</i>	14 \pm 5.35	p = 0.0401
Figure 4B					
N=3	Q40 on control RNAi	34 \pm 1.53	Q40; <i>lir-3(pk2183)</i> on control RNAi	15 \pm 0.99	< 0.0001
	Q40 on <i>hoe-1 (RNase Z)</i>	37 \pm 1.11	Q40; <i>lir-3(pk2183)</i> on <i>hoe-1 (RNase Z)</i>	17 \pm 0.94	< 0.0001
	Q40 on <i>hpo-31</i>	36 \pm 2.14	Q40; <i>lir-3(pk2183)</i> on <i>hpo-31</i>	14 \pm 0.70	< 0.0001
	Q40 on <i>T06D8.9 (TSEN15)</i>	38 \pm 1.82	Q40; <i>lir-3(pk2183)</i> on <i>T06D8.9 (TSEN15)</i>	16 \pm 1.41	< 0.0001
	Q40 on <i>F52C12.3 (TSEN54)</i>	33 \pm 1.30	Q40; <i>lir-3(pk2183)</i> on <i>F52C12.3 (TSEN54)</i>	16 \pm 1.13	< 0.0001
	Q40 on <i>rtcb-1</i>	29 \pm 1.47	Q40; <i>lir-3(pk2183)</i> on <i>rtcb-1</i>	16 \pm 0.70	< 0.0001
	Q40 on <i>xpo-3</i>	39 \pm 1.59	Q40; <i>lir-3(pk2183)</i> on <i>xpo-3</i>	16 \pm 0.99	< 0.0001
	Q40 on <i>pus-1</i>	30 \pm 1.58	Q40; <i>lir-3(pk2183)</i> on <i>pus-1</i>	16 \pm 0.79	< 0.0001
	Q40 on <i>clpf-1</i>	38 \pm 1.01	Q40; <i>lir-3(pk2183)</i> on <i>clpf-1</i>	17 \pm 1.37	< 0.0001
	Q40 on <i>T04H1.2 (GTBP2)</i>	31 \pm 1.16	Q40; <i>lir-3(pk2183)</i> on <i>T04H1.2 (GTBP2)</i>	16 \pm 0.86	< 0.0001
Figure S4A					
N=4	Q40 on control RNAi	26 \pm 4.81	Q40; <i>lir-3(tm813)</i> on control RNAi	8 \pm 1.34	p = 0.0055
	Q40 on <i>ama-1 (Pol II)</i>	20 \pm 1.12	Q40; <i>lir-3(tm813)</i> on <i>ama-1 (Pol II)</i>	8 \pm 0.88	p = 0.0001
	Q40 on <i>rpc-1 (Pol III)</i>	24 \pm 2.77	Q40; <i>lir-3(tm813)</i> on <i>rpc-1 (Pol III)</i>	7 \pm 1.03	p = 0.0006
	Q40 on <i>tbp-1 (TBP)</i>	26 \pm 2.91	Q40; <i>lir-3(tm813)</i> on <i>tbp-1 (TBP)</i>	9 \pm 1.86	p = 0.0013
	Q40 on <i>tftc-5 (TFC-1)</i>	29 \pm 3.38	Q40; <i>lir-3(tm813)</i> on <i>tftc-5 (TFC-1)</i>	11 \pm 2.16	p = 0.0021
	Q40 on <i>npp-13</i>	24 \pm 2.25	Q40; <i>lir-3(tm813)</i> on <i>npp-13</i>	8 \pm 1.19	p = 0.0004
Figure S4B					
N=3	Q40 on control RNAi	23 \pm 1.71	Q40; <i>lir-3(pk2183)</i> on control RNAi	11 \pm 1.15	< 0.0001
	Q40 on <i>cars-1</i>	19 \pm 1.42	Q40; <i>lir-3(pk2183)</i> on <i>cars-1</i>	9 \pm 0.66	< 0.0001
	Q40 on <i>frs-1</i>	20 \pm 1.60	Q40; <i>lir-3(pk2183)</i> on <i>frs-1</i>	11 \pm 0.74	< 0.0001
	Q40 on <i>nrs-1</i>	24 \pm 1.70	Q40; <i>lir-3(pk2183)</i> on <i>nrs-1</i>	14 \pm 1.28	< 0.0001
	Q40 on <i>krs-1</i>	25 \pm 1.91	Q40; <i>lir-3(pk2183)</i> on <i>krs-1</i>	9 \pm 1.00	< 0.0001
	Q40 on <i>aars-1</i>	27 \pm 1.88	Q40; <i>lir-3(pk2183)</i> on <i>aars-1</i>	14 \pm 2.11	< 0.0001
	Q40 on <i>trs-1</i>	20 \pm 1.41	Q40; <i>lir-3(pk2183)</i> on <i>trs-1</i>	13 \pm 1.04	< 0.0001
	Q40 on <i>wrs-2</i>	23 \pm 1.06	Q40; <i>lir-3(pk2183)</i> on <i>wrs-2</i>	13 \pm 1.47	< 0.0001
	Q40 on <i>rrt-1</i>	20 \pm 1.44	Q40; <i>lir-3(pk2183)</i> on <i>rrt-1</i>	10 \pm 0.79	< 0.0001
	Q40 on <i>lrs-1</i>	20 \pm 1.28	Q40; <i>lir-3(pk2183)</i> on <i>lrs-1</i>	15 \pm 1.61	p = 0.0179

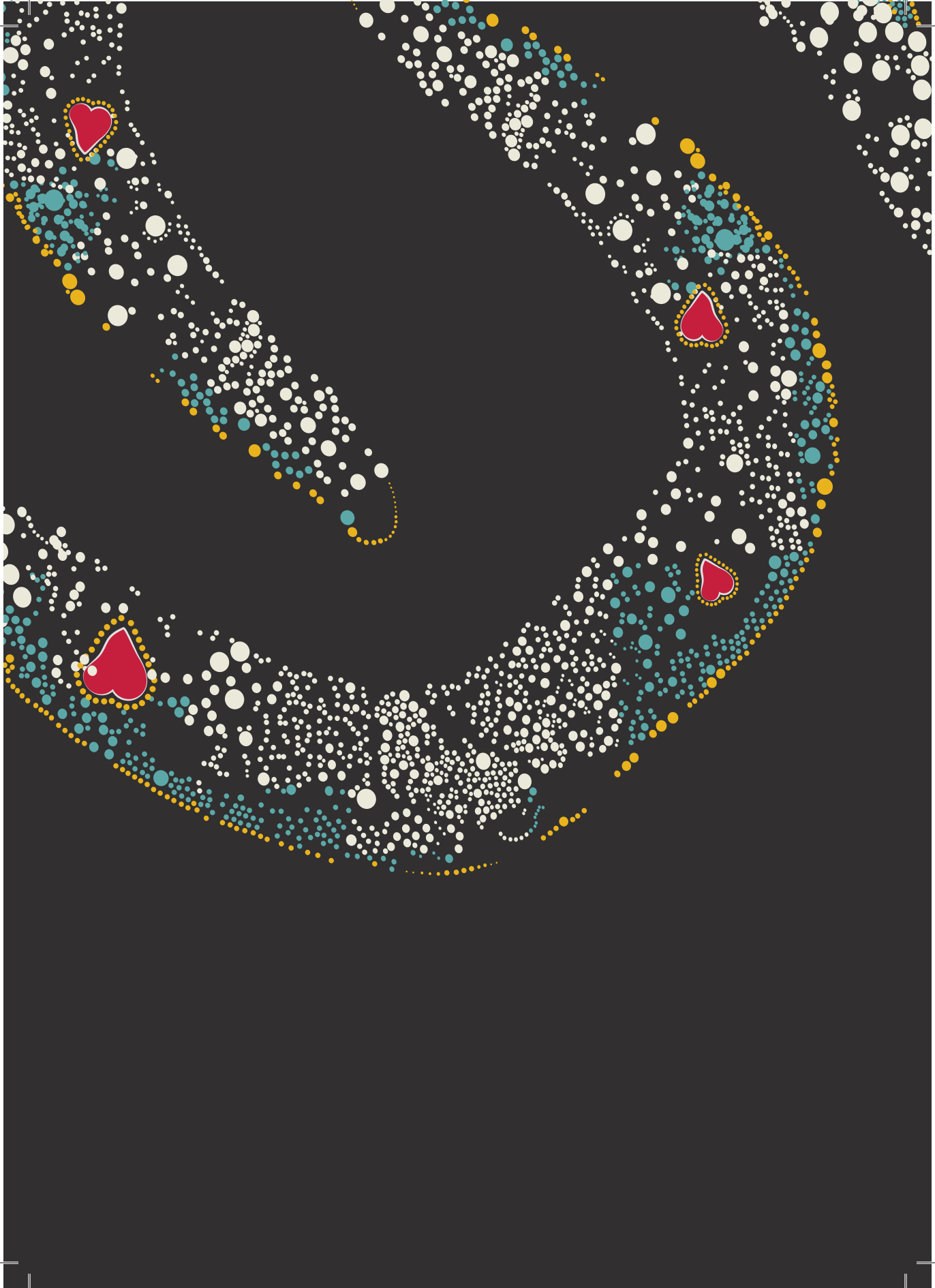
Chapter IV

<i>Q40 on grs-1</i>	15 ± 0.92	<i>Q40;lir-3(pk2183) on grs-1</i>	8 ± 0.60	< 0.0001
<i>Q40 on prs-1</i>	20 ± 1.27	<i>Q40;lir-3(pk2183) on prs-1</i>	10 ± 0.84	< 0.0001
<i>Q40 on dars-1</i>	17 ± 1.34	<i>Q40;lir-3(pk2183) on dars-1</i>	8 ± 0.82	< 0.0001
<i>Q40 on irs-1</i>	19 ± 1.29	<i>Q40;lir-3(pk2183) on irs-1</i>	10 ± 1.03	< 0.0001
<i>Q40 on qars-1</i>	22 ± 1.59	<i>Q40;lir-3(pk2183) on qars-1</i>	10 ± 0.73	< 0.0001

ns - not significant

Table S2. List of RNAi foods used in knockdown experiments.

Sequence name	Gene	Description
L4440	-	Empty vector
F36A4.7	ama-1	Large subunit of RNA Polymerase II
C42D4.8	rpc-1	Ortholog of the human RNA Polymerase III
T20B12.2	tbp-1	Ortholog of the human TATA-box-binding protein (TBP)
T02C12.3	tftc-5	Ortholog of the human general transcription factor IIIC (TFC-1)
Y37E3.15	npp-3	Nucleoporin required for normal nuclear pore complex distribution in the nuclear envelope; also described as being part of the RNA Polymerase III complex
E04A4.4	hoe-1	RNAse Z
F55B12.4	hpo-31	Predicted to have tRNA nucleotidyl transferase activity
T06D8.9	-	43% homologous to the human TSEN15 subunit of the tRNA splicing endonuclease
F52C12.3	-	31% homologous to the human TSEN54 subunit of the tRNA splicing endonuclease
F16A11.2	rtcb-1	Ortholog of the human tRNA splicing ligase (RTCB)
C49H3.10	xpo-3	Ortholog of the human importin beta 1
W06H3.2	pus-1	Putative mitochondrial tRNA synthetase
F29A2.4	clpf-1	48% homologous to mammalian RNA kinase protein, recently shown to associate with TSEN and drive tRNA splicing
T04H1.2	-	45% homologous to the human GTB-binding protein 2 protein, recently shown to prevent ribosome stalling
Y23H5A.7	cars-1	Cysteinyl amino-acyl tRNA synthetase
T08B2.9	frs-1	Phenylalanyl amino-acyl tRNA synthetase
F22D6.3	nrs-1	Asparaginyl amino-acyl tRNA synthetase
T02G5.9	krs-1	Lysyl amino-acyl tRNA synthetase
W02B12.6	aars-1	Alanyl amino-acyl tRNA synthetase
C47D12.6	trs-1	Threonyl amino-acyl tRNA synthetase
C34E10.4	wrs-2	Tryptophanyl amino-acyl tRNA synthetase
F26F4.10	rrt-1	Arginyl amino-acyl tRNA synthetase
R74.1	lrs-1	Leucyl amino-acyl tRNA Synthetase
T10F2.1	grs-1	Glycyl amino-acyl tRNA synthetase
T20H4.3	prs-1	Prolyl amino-acyl tRNA synthetase
B0464.1	dars-1	Aspartyl amino-acyl tRNA synthetase
R11A8.6	irs-1	Isoleucyl amino-acyl tRNA synthetase
Y41E3.4	qars-1	Glutaminyl amino-acyl tRNA synthetase





CHAPTER V

Transcriptional profiling in *C. elegans* reveals developmental and cellular responses to polyglutamine expansion proteins



Olga Sin^{1,2}, Tristan van de Jong^{1*}, Renée Seinstra¹,
Ellen A. A. Nollen¹, Victor Guryev¹*

** equal contribution*

¹ European Research Institute for the Biology of Aging, University of Groningen,
University Medical Centre Groningen, Groningen, The Netherlands

² Graduate Program in Areas of Basic and Applied Biology, Abel Salazar Biomedical
Sciences Institute, University of Porto, Porto, Portugal

Work in progress

Abstract

Protein aggregation is a hallmark of neurodegenerative diseases such as Alzheimer's, Parkinson's and polyglutamine diseases. Polyglutamine diseases are characterized by the abnormal expansion of CAG repeats in the disease-causing gene. Consequently, when the gene is translated into a protein, the protein is unable to acquire its native conformation, so that it becomes misfolded and aggregation-prone. The aggregation-prone protein interacts aberrantly with other normal, functioning proteins in the cell, ultimately disrupting protein homeostasis. While most research into polyglutamine diseases has focused on the changes that take place after the aggregation-prone protein has been translated, less is known about changes at the transcriptional level. We therefore sought to identify the transcriptional changes that occur when a cell is exposed to aggregation-prone proteins such as polyglutamine expansion proteins. We took advantage of a *C. elegans* model for polyglutamine diseases – which recapitulates the protein misfolding and aggregation observed in human disease – and performed whole transcriptome analysis at two different stages of larval development in both wild type (N2) and polyglutamine-expressing animals.

Principal component analysis revealed that the major sources of expression variation were related to stage of development (L2/L3 versus L4) followed by the genetic background (wild type N2 versus the polyglutamine model). RNA sequencing-derived data was analyzed by combining standard differential expression analysis with gene ontology enrichment analysis to identify those genes whose expression was affected by aggregation-prone proteins. The genes thus identified were found to be related to the unfolded protein response, oxidative stress and immune response. We also found that aggregation-prone proteins slow down the rate of animal development. We propose that this combination of events may represent a coping strategy that the organism adopts to preserve its fitness as a response to the pathological presence of aggregation-prone proteins.

Introduction

Polyglutamine diseases are a class of neurodegenerative disorders that include Huntington's disease, spinocerebellar ataxia (types 1, 2, 6, 7, 17), Machado–Joseph disease and spinobulbar muscular atrophy [1, 2]. Due to a loss of neurons, affected patients commonly experience motor dysfunction as well as cognitive and behavioral abnormalities [3]. These diseases are characterized by an abnormal expansion of CAG repeats in the disease-causing genes. The resulting protein has an unusually long stretch of glutamines, which causes it to become misfolded and prone to aggregation, ultimately contributing to protein aggregation [4]. Despite our knowledge of the underlying causes of polyglutamine diseases, we are not yet able to explain why protein aggregation occurs in affected patients. If we can understand how the cell copes with protein aggregation, we can identify and characterize the cellular pathways implicated in pathogenesis and target these for disease-modifying treatments.

Polyglutamine aggregation has been recapitulated in several models in the nematode *Caenorhabditis elegans* [5–8]. In one of those models, an expanded polyglutamine stretch containing 40 CAG repeats is expressed in the body wall muscle cells, whereby the formation of aggregates is proteotoxic and progresses with age [6]. Here we used this model to find answers to the following questions: What transcriptional changes occur in the cell upon expression of polyglutamine expansion proteins? Which pathways are activated? Does this change during the course of worm development? To this end, we performed a global integrative analysis of the transcriptome of a *C. elegans* model of polyglutamine diseases, by employing principal component analysis and differential gene expression analysis. We aimed to identify the pathways that are activated upon pathological overexpression of polyglutamine expansion proteins in order to gain insight into the relevant cellular responses.

Results

Developmental stage and genetic background account for differences in transcriptome profiles

In our analysis of the transcriptome of polyglutamine-expressing (Q40) animals, we selected larval stages two/three (L2/L3) and four (L4), in order to examine transcriptome changes during development. To search for the main sources of variance in our transcriptome profiling data, we employed principal component analysis (PCA). The first principal component yielded a clear distinction between L2/L3 and L4 animals, showing that the developmental stage had a strong effect in our expression data (Figure 1A, C). The following principal component explained the variance between our datasets based on the presence of the transgene (Figure 1B, C). Indeed, while the profile of polyglutamine-expressing animals was very similar to that of wild type N2 animals at the L2/L3 stage, these differences were more pronounced at developmental stage L4 (Figure 1A-C). These results are consistent with the observation that – under our laboratory conditions – expression of polyglutamine expansion proteins slows down worm development (Figure 1D). We performed enrichment analysis based on gene ontology (GO) using the web-based toolset gProfiler (<http://biit.cs.ut.ee/gprofiler/>, Ensembl 79). This enabled us to learn more about the gene set comprising the first principal component, since this component explained most of the variance of the data (18.7%). GO term analysis of the first principal component revealed that the genes in the first principal component were related to developmental processes (Figure 1E). In all, our transcriptomic profiling data demonstrated that expression variation was determined firstly by developmental stage and secondly by genetic background.

Polyglutamine expansion proteins trigger the ER-associated unfolded protein response

We next used differential gene expression analysis to look for genes that were significantly up- or downregulated due to the presence of aggregation-prone proteins. A false-discovery rate (FDR) cutoff of 0.1% was used to identify such differentially expressed genes. First, we excluded genes differentially expressed between the L2/L3 and L4 stages in the N2 genetic background, thus avoiding developmental effects (and therefore false positives) that might

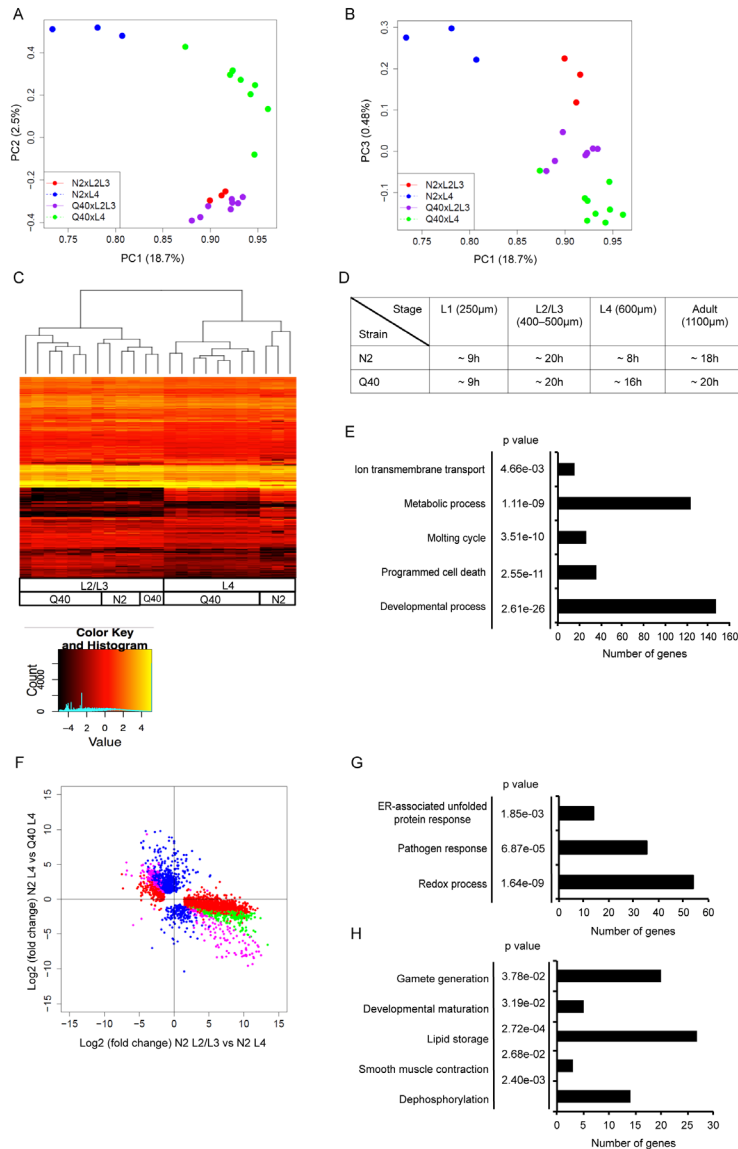


Figure 1. Effect of aggregation-prone proteins on the *C. elegans* transcriptome. (A, B) PCA plots of the 23 samples included in this study. Different colors indicate the different time points (L2/L3 or L4) and genetic backgrounds (Q40 or N2). PC, principal component. **(C)** Heat map showing the distribution of the 23 samples included in this study. The hierarchical clustering was generated using the average linkage cluster method with a Euclidean metric distance. **(D)** Approximate number of hours from egg hatching until L1 and between successive developmental stages until adulthood in *C. elegans*. **(E)** Gene ontology (GO) term enrichment analysis for gene set of the first principal component. **(F)** Results of differential gene expression analysis as measured by the fold change in expression between raw reads of wild type animals in L2/L3 and L4 stage (x axis) and N2 and Q40 animals in L4 stage (y axis). **(G)** GO term enrichment analysis for genes differentially expressed in the presence of aggregation-prone proteins (blue data points from panel (F)), and the number of genes associated with each GO term. **(H)** GO term enrichment analysis for genes differentially expressed in the presence of aggregation-prone proteins (green data points from panel (F)), and the number of genes associated with each term.

mask the contribution of the polyglutamine expansion proteins to the cell's transcriptome (Figure 1F, red). Then, to determine the specific contribution of polyglutamine expansion proteins to the cell's transcriptome, we looked for genes that were differentially expressed between the Q40 and N2 animals at the L4 stage (Figure 1F, blue). GO term enrichment analysis of this cluster identified genes involved in three types of cellular functions, namely the unfolded protein response (UPR, 14 genes), the pathogen response (35 genes), and the redox process (54 genes) (Figure 1G). Further inspection of the genes involved in the UPR pathway revealed significant downregulation of the *abu* and *pqn* family of genes (Table 1). The *abu* (activated in blocked UPR) family of genes is expressed when the UPR is blocked and it has been suggested that their protein products may protect the cell against aberrantly folded proteins [9]. Both the *pqn* and *abu* genes encode proteins whose amino acid sequences contain glutamine/asparagine-rich (Q/N) domains, a common feature of prion-like domains [10, 11]. This hints at the possibility that polyglutamine expansion proteins trigger the UPR to cope with misfolded proteins in the cell, thereby downregulating the *abu/pqn* genes since they are not necessary.

Parallel to this, in Q40 animals we detected upregulation of the *cdr-4*, *gad-3* and *C14B9.2* genes (Table 1), all of which are regulated by *xbp-1*, a transcription factor required for the activation of the UPR [9, 12, 13]. While not much is known about *cdr-4* and *gad-3*, *C14B9.2* is known to encode a disulfide isomerase. Protein disulfide isomerases are important for the formation of native disulfide bonds in proteins transiting through the ER, a process known as oxidative protein folding (reviewed in [14]). We observed no activation of other canonical UPR components, such as the worm homologs of BiP/GRP78 (*hsp-3* and *hsp-4*), an ER-associated molecular chaperone [15-17]. Overall, this set of differentially expressed genes is consistent with the cellular response pathways known to be triggered by misfolded, aggregation-prone proteins in the ER [18].

Curiously, GO term enrichment analysis also identified a cluster of 35 genes involved in the pathogen response, including *lys-7*, *sdd-3* and *cnc-6* (Table 1). Aggregation-prone proteins – and in particular polyglutamine expansion proteins – are not known to trigger a pathogen defense response. Moreover, several genes within this cluster are expressed in the *C. elegans* gut, suggesting

that aggregation-prone proteins can also provoke cell non-autonomous effects. Finally, we identified upregulation of genes within the redox cluster, including genes known to be involved in oxidative stress (*cst-2* and *sod-3*) and belonging to the cytochrome P450 family of monooxygenases, including *kmo-1* (Table 1). The yeast homolog of *kmo-1* – *bna-4* – has been identified as a potent suppressor of polyglutamine-induced proteotoxicity [19]. Taken together, these results point to aggregation-prone proteins inducing cellular stress, resulting in activation of cellular stress response pathways.

Table 1. Genes expressed differentially between wild type and polyglutamine-expressing animals at the L4 stage (blue data points from Figure 1F).

Cellular process (no of genes)	Function (no of genes)	Sequence name	Gene examples	Fold change	Short description
Response to stress (53)	Unfolded protein response (14)	AC3.3	abu-1	-2,66	Transmembrane proteins with a glutamine/asparagine-rich ("prion") domain
		C03A7.7	abu-6	-2,35	
		C03A7.8	abu-7	-2,28	
		R09F10.2	abu-9	-2,70	
		F35A5.3	abu-10	-1,99	
		T01D1.6	abu-11	-2,16	Proteins with a glutamine/asparagine-rich ("prion") domain
		C03A7.4	abu-15	-2,26	
		AC3.4	pqn-2	-2,70	
		R09F10.7	pqn-57	-3,15	
		C14B9.2	-	1,40	
	Pathogen response (35)	K01D12.11	cdr-4	1,79	Predicted transmembrane protein part of the cadmium responsive (CDR) family of proteins
		B0222.9	gad-3	1,36	Ortholog of human xanthine dehydrogenase
		C02A12.4	lys-7	2,00	Antimicrobial enzyme
		T08A9.7	spp-3	1,17	Antimicrobial peptide
Metabolic process (91)	Redox process (54)	Y46E12A.1	cnc-6	7,42	Antimicrobial peptide
		C24A8.4	cst-2	1,00	Protein kinase
		C08A9.1	sod-3	2,54	Iron/manganese superoxide dismutase
		R07B7.5	kmo-1	1,67	Mitochondrial enzyme
		F02C12.5	cyp-13B1	5,75	Membrane-associated, heme-containing NADPH-dependent monooxygenases
		K09A11.3	cyp-14A2	1,58	
		R04D3.1	cyp-14A4	3,24	
		T10B9.7	cyp-13A2	1,77	
		B0304.3	cyp-23A1	2,45	

Polyglutamine expansion proteins affect *C. elegans* development

While performing differential expression analysis (also using an FDR cutoff of 0.1%), we identified additional clusters of differentially expressed genes (Figure 1F, magenta and green). These clusters of genes were significantly upregulated during development of wild type animals while slightly downregulated in Q40 animals. Further inspection of this cluster identified two sub-groups: one sub-group had more genes that were strongly downregulated (Figure 1F, magenta)

while the other had more genes that were clustered but less downregulated (Figure 1F, green). Since we suspected that this expression pattern may be due to developmental differences provoked by the expression of aggregation-prone proteins, we performed GO term analysis of each cluster individually. For the sub-group with strongly downregulated genes, we could draw no conclusions regarding biological processes (Figure 1F, magenta). However, the clustered genes (Figure 1F, green) were enriched for development and maturation processes, muscle reorganization and lipid storage (Figure 1H). The fact that we observed downregulation of developmental genes as a response to the presence of aggregation-prone proteins is consistent with the aforementioned developmental delay observed in polyglutamine-expressing animals (Figure 1D), suggesting that aggregation-prone proteins can also affect animal development. In all, expression of aggregation-prone proteins alters the expression developmental genes, which slows down *C. elegans* development.

Discussion

Here we investigated the cellular responses that are elicited by the presence of aggregation-prone proteins using a *C. elegans* model of polyglutamine diseases. We used differential gene expression analysis to demonstrate that the presence of aggregation-prone proteins – in the form of polyglutamine expansion proteins – triggers the expression of ER-associated UPR genes. This observation is unexpected since polyglutamine expansion proteins do not usually enter the secretory pathway. We speculate that this cellular stress response is activated in an attempt to reduce overall protein synthesis of misfolded polyglutamine expansion proteins and perhaps that of other polyglutamine-rich proteins – thereby explaining the downregulation of *abu/pqn* genes – as has been suggested by others [18].

The presence of polyglutamine expansion proteins is thought to induce the heat shock response (HSR), a cellular stress response that relies on molecular chaperones to prevent protein misfolding and aggregation [20, 21]. We did not, however, detect any differential expression of genes associated with the HSR. One possibility is that the animals in our experiment have undergone adaptive evolution, whereby over successive generations of growth in laboratory conditions the animals have developed the ability to cope with the constitutive

expression of polyglutamine expansion proteins, while maintaining the aggregation-associated phenotypes. Such adaptive evolution has for example occurred in the wild type strain N2, which after continuous maintenance in laboratory conditions has developed “laboratory-derived” alleles as an adaptation process ([22, 23] also reviewed in [24]).

The fact that we saw upregulation of several genes involved in the innate immune response suggests that polyglutamine expansion proteins also affect immune response signaling pathways. The induction of a defense response has not been previously described in polyglutamine diseases. We propose that the presence of polyglutamine expansion proteins is perceived by the worms as an infection, thereby triggering the innate immune system. This was recently shown to be the case for the aggregation-prone proteins TDP-43 and FUS, which induce immune responses in *C. elegans* models of amyotrophic lateral sclerosis [25]. It is of interest to note that the UPR is also part of the cellular defense in *C. elegans* and is activated upon pathogen infection, further supporting the notion that polyglutamine expansion proteins induce effects in the organism similar to those induced by pathogens [26-28].

Our findings relating to the induction of cellular stress pathways are not entirely without precedent. Oxidative stress originates from the imbalance between the production and clearance of reactive oxygen species (ROS) and has been linked to Huntington’s disease [29, 30]. Here we identified several genes with redox functions, which points to another cellular stress response activated in the presence of polyglutamine expansion proteins. The fact that we observed the involvement of the UPR, the immune response and oxidative stress pathways suggests that aggregation-prone proteins have a systemic effect in the organism. It will be interesting to further address how these cellular pathways are functionally affected in mammalian neuronal cell or animal models and determine whether these pathways mediate neurodegeneration induced by polyglutamine expansion proteins in these systems.

Finally, we found larval development to be significantly affected in polyglutamine-expressing animals. Others have reported that expression of aggregation-prone proteins slows down the growth rate of yeast [18] and delays

development in another *C. elegans* model expressing polyglutamine expansion proteins in the body wall muscle cells [21]. Collectively, these findings point towards the tentative notion of a delay in organismal development as a result of cell non-autonomous signaling across tissues that is provoked by the presence of aggregation-prone proteins.

In conclusion, we propose that the presence of aggregation-prone proteins activates a cellular stress response that – combined with a delay in development – may impose a fitness cost in this model.

Methods

Strains and genetics

Standard methods were used for culturing *C. elegans* at 20°C [31]. To synchronize animals, eggs were collected from gravid hermaphrodites by hypochlorite bleaching and hatched overnight in M9 buffer. The desired number of L1 animals were subsequently cultured on agar plates (#633185, Greiner Bio-One) containing nematode growth medium and seeded with OP50 bacteria. The following strains were used or generated: N2 (wild type): AM141 *rmls133[P(unc-54)Q40::YFP]X* [6]; OW1004 *rmls133[P(unc-54)Q40::YFP]X*; and OW1021 *rmls133[P(unc-54)Q40::YFP]X*.

RNA sequencing

Worms were grown to larval stages two/three (L2/3) and four (L4) and total RNA was extracted using TRIzol Reagent (Life Technologies) according to the manufacturer's instructions. For polyA RNA sequencing, sequencing libraries were prepared using the TruSeq Sample Preparation V2 Kit (Illumina) and subjected to high-throughput single-end sequencing (50 bp) in an Illumina HiSeq 2500 instrument.

Data analysis

Single-end reads obtained by Illumina sequencing were aligned to the *C. elegans* reference genome (Wormbase, version WB235) using Star Aligner [32]. Raw read counts were extracted from the alignments using the HTSeq package. The reads were fragment per million (FPM) normalized, log transformed and null averaged prior to Principal Component Analysis (PCA). Average clustering

was performed using Pearson correlation for the corrected RNA samples and complete clustering was performed with euclidean distances for the read distance, this was plotted with the heatmap2 function in R. Differential expression analysis was performed on the corrected raw reads using EdgeR [33], and gene ontology enrichment analysis was performed using the web-based toolset gProfiler [34, 35].

Acknowledgements

Alumni chapter Gooische Groningers facilitated by the Ubbo Emmius Fund. We thank Sally Hill for editing the manuscript. This work was supported by the European Research Council (ERC) starting grant (to E.A.A.N.), the Fundação para a Ciência e Tecnologia (SFRH/BPD/51009/2010) (to O. S.).

Author contributions

O.S. and E.A.A.N. designed the experiments. O.S. and R.S. performed the experiments. O.S., T.v.d.J. and V.G. analyzed the data. O.S., E.A.A.N. and V.G. wrote the manuscript.

References

1. Blum ES, Schwendeman AR, Shaham S (2013) PolyQ disease: misfiring of a developmental cell death program? *Trends in Cell Biology* 23:168–174. doi: 10.1016/j.tcb.2012.11.003
2. La Spada AR, Taylor JP (2010) Repeat expansion disease: progress and puzzles in disease pathogenesis. *Nature Reviews Genetics* 11:247–258. doi: 10.1038/nrg2748
3. Orr HT, Zoghbi HY (2007) Trinucleotide repeat disorders. *Annu Rev Neurosci* 30:575–621. doi: 10.1146/annurev.neuro.29.051605.113042
4. Taylor JP, Hardy J, Fischbeck KH (2002) Toxic proteins in neurodegenerative disease. *Science* 296:1991–1995. doi: 10.1126/science.1067122
5. Brignull HR, Moore FE, Tang SJ, Morimoto RI (2006) Polyglutamine proteins at the pathogenic threshold display neuron-specific aggregation in a pan-neuronal *Caenorhabditis elegans* model. *Journal of Neuroscience* 26:7597–7606. doi: 10.1523/JNEUROSCI.0990-06.2006
6. Morley JF, Brignull HR, Weyers JJ, Morimoto RI (2002) The threshold for polyglutamine-expansion protein aggregation and cellular toxicity is dynamic and influenced by aging in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences* 99:10417–10422. doi: 10.1073/pnas.152161099
7. Parker JA, Connolly JB, Wellington C, et al. (2001) Expanded polyglutamines in *Caenorhabditis elegans* cause axonal abnormalities and severe dysfunction of PLM mechanosensory neurons without cell death. *Proceedings of the National Academy of Sciences* 98:13318–13323. doi: 10.1073/pnas.231476398
8. Faber PW, Alter JR, MacDonald ME, Hart AC (1999) Polyglutamine-mediated dysfunction and apoptotic death of a *Caenorhabditis elegans* sensory neuron. *Proceedings of the National Academy of Sciences* 96:179–184.
9. Urano F, Calton M, Yoneda T, et al. (2002) A survival pathway for *Caenorhabditis elegans* with a blocked unfolded protein response. *J Cell Biol* 158:639–646. doi: 10.1083/jcb.200203086
10. Maddelein ML, Wickner RB (1999) Two prion-inducing regions of Ure2p are nonoverlapping. *Molecular and Cellular Biology* 19:4516–4524.
11. DePace AH, Santoso A, Hillner P, Weissman JS (1998) A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion. *Cell* 93:1241–1252.
12. Calton M, Zeng H, Urano F, et al. (2002) IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA. *Nature* 415:92–96. doi: 10.1038/415092a
13. Shen X, Ellis RE, Lee K, et al. (2001) Complementary signaling pathways regulate the unfolded protein response and are required for *C. elegans* development. *Cell* 107:893–903.
14. Riemer J, Bulleid N, Herrmann JM (2009) Disulfide formation in the ER and mitochondria: two solutions to a common process. *Science* 324:1284–1287. doi: 10.1126/science.1170653
15. Nishikawa SI, Fewell SW, Kato Y, et al. (2001) Molecular chaperones in the yeast endoplasmic reticulum maintain the solubility of proteins for retrotranslocation and degradation. *J Cell Biol* 153:1061–1070.
16. Schmitz A, Maintz M, Kehle T, Herzog V (1995) In vivo iodination of a misfolded proinsulin reveals co-localized signals for Bip binding and for degradation in the ER. *EMBO J* 14:1091–1098.
17. Knittler MR, Dirks S, Haas IG (1995) Molecular chaperones involved in protein degradation in the endoplasmic reticulum: quantitative interaction of the heat shock cognate protein BiP with partially folded immunoglobulin light chains that are degraded in the endoplasmic reticulum. *Proceedings of the National Academy of Sciences* 92:1764–1768.
18. Geiler-Samerotte KA, Dion MF, Budnik BA, et al. (2011) Misfolded proteins impose a dosage-dependent fitness cost and trigger a cytosolic unfolded protein response in yeast. *Proc Natl Acad Sci USA* 108:680–685. doi: 10.1073/pnas.1017570108
19. Giorgini F, Guidetti P, Nguyen Q, et al. (2005) A genomic screen in yeast implicates kynurenine 3-monooxygenase as a therapeutic target for Huntington disease. *Nat Genet* 37:526–531. doi: 10.1038/ng1542
20. Morley JF, Morimoto RI (2004) Regulation of longevity in *Caenorhabditis elegans* by heat shock factor and molecular chaperones. *Mol Biol Cell* 15:657–664. doi: 10.1091/mbc.E03-07-0532
21. Satyal SH, Schmidt E, Kitagawa K, et al. (2000) Polyglutamine aggregates alter protein folding homeostasis in *Caenorhabditis elegans*.

- Proceedings of the National Academy of Sciences 97:5750–5755. doi: 10.1073/pnas.100107297
22. Duveau F, Félix M-A (2012) Role of pleiotropy in the evolution of a cryptic developmental variation in *Caenorhabditis elegans*. PLoS Biol 10:e1001230. doi: 10.1371/journal.pbio.1001230
23. Weber KP, De S, Kozarewa I, et al. (2010) Whole genome sequencing highlights genetic changes associated with laboratory domestication of *C. elegans*. PLoS ONE 5:e13922. doi: 10.1371/journal.pone.0013922
24. Sterken MG, Snoek LB, Kammenga JE, Andersen EC (2015) The laboratory domestication of *Caenorhabditis elegans*. Trends in Genetics 31:224–231. doi: 10.1016/j.tig.2015.02.009
25. Vérièpe J, Fossouo L, Parker JA (2015) Neurodegeneration in *C. elegans* models of ALS requires TIR-1/Sarm1 immune pathway activation in neurons. Nat Commun 6:7319. doi: 10.1038/ncomms8319
26. Richardson CE, Kooistra T, Kim DH (2010) An essential role for XBP-1 in host protection against immune activation in *C. elegans*. Nature 463:1092–1095. doi: 10.1038/nature08762
27. Bischof LJ, Kao C-Y, Los FCO, et al. (2008) Activation of the unfolded protein response is required for defenses against bacterial pore-forming toxin in vivo. PLoS Pathog 4:e1000176. doi: 10.1371/journal.ppat.1000176
28. Haskins KA, Russell JF, Gaddis N, et al. (2008) Unfolded protein response genes regulated by CED-1 are required for *Caenorhabditis elegans* innate immunity. Developmental Cell 15:87–97. doi: 10.1016/j.devcel.2008.05.006
29. Melkani GC, Trujillo AS, Ramos R, et al. (2013) Huntington's disease induced cardiac amyloidosis is reversed by modulating protein folding and oxidative stress pathways in the *Drosophila* heart. PLoS Genet 9:e1004024. doi: 10.1371/journal.pgen.1004024
30. Wyttenbach A, Sauvageot O, Carmichael J, et al. (2002) Heat shock protein 27 prevents cellular polyglutamine toxicity and suppresses the increase of reactive oxygen species caused by huntingtin. Human Molecular Genetics 11:1137–1151.
31. Brenner S (1974) The genetics of *Caenorhabditis elegans*. Genetics 77:71–94.
32. Dobin A, Davis CA, Schlesinger F, et al. (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21. doi: 10.1093/bioinformatics/bts635
33. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140. doi: 10.1093/bioinformatics/btp616
34. Reimand J, Arak T, Vilo J (2011) g:Profiler—a web server for functional interpretation of gene lists (2011 update). Nucleic Acids Research 39:W307–15. doi: 10.1093/nar/gkr378
35. Reimand J, Kull M, Peterson H, et al. (2007) g:Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. Nucleic Acids Research 35:W193–200. doi: 10.1093/nar/gkm226

Supplemental Information

Table S1. Top 100 genes upregulated in polyglutamine-expressing animals at the L4 stage (blue data points from Figure 1F).

Sequence name	Gene	False discovery rate (FDR)	Fold change
F01D4.8		4,90E-28	5,33
W02H5.6	srh-271	1,16E-24	6,41
C32C4.7	ugt-35	3,06E-22	9,80
F02C12.5	cyp-13B1	1,14E-21	5,75
Y9C9A.16		6,40E-20	5,80
C33H5.1		9,09E-19	8,87
B0303.14		9,23E-19	3,31
T02B11.4		8,19E-18	2,76
F10B5.3		4,24E-17	4,37
K03H6.2		4,24E-17	3,26
F15B9.1	far-3	4,19E-15	5,26
K09D9.9		4,77E-15	8,34
C52A10.2		1,16E-14	4,33
F49C12.2		2,60E-14	8,37
K08C7.5	fmo-2	3,08E-14	5,38
F26A1.8		3,15E-14	5,20
H22D14.1	nhr-267	7,95E-14	5,73
ZC373.7	col-176	1,29E-13	6,58
D2023.7	col-158	1,63E-13	5,45
F27C8.4	spp-18	4,80E-13	2,73
F49C12.4		5,03E-13	8,81
B0213.6	nlp-31	5,62E-13	2,73
C02A12.1	gst-33	6,50E-13	3,15
F25D1.5		8,49E-13	2,86
Y51A2D.5	hmit-1.2	1,16E-12	2,09
K07C6.4	cyp-35B1	1,31E-12	4,36
Y80D3A.7	ptr-22	3,43E-12	4,98
C33H5.2		4,67E-12	7,73
F48C1.9		4,68E-12	7,44
B0238.12		5,37E-12	2,99
Y38E10A.15	nspe-7	6,50E-12	3,85
K11G9.6	mtl-1	6,99E-12	8,32
Y45F10D.15		9,69E-12	7,29
F53A9.8		1,03E-11	2,66
C50B6.7		1,49E-11	2,12
B0238.1		1,56E-11	3,10
Y48E1B.10	gst-20	1,75E-11	2,85
T03F7.3	srh-46	1,81E-11	6,20
Y71G12B.6		1,86E-11	2,40
M162.5		1,97E-11	2,24
Y54G2A.37		3,26E-11	3,47
D1086.3		4,20E-11	3,58
F49C12.5		4,99E-11	7,28
T27D12.6		5,28E-11	4,83
F14A5.1	nhr-264	6,08E-11	5,49
B0511.1	fkf-7	8,99E-11	3,38
W06H12.1	ztf-6	1,09E-10	2,26
H41C03.1		1,15E-10	3,08
E02H1.7	nhr-19	1,65E-10	2,17
ZK1067.6	sym-2	2,49E-10	2,49
C40D2.1	math-19	3,32E-10	4,27
T22A3.8	lam-3	3,71E-10	2,48
K07C6.3	cyp-35B2	3,78E-10	6,92
M7.3	bcc-1	3,85E-10	2,28
F09C8.1		3,88E-10	3,29
F54F3.3	lipl-1	4,26E-10	3,17
Y49E10.18		4,87E-10	1,78

Transcriptional profiling in *C. elegans* reveals developmental and cellular responses
to polyglutamine expansion proteins

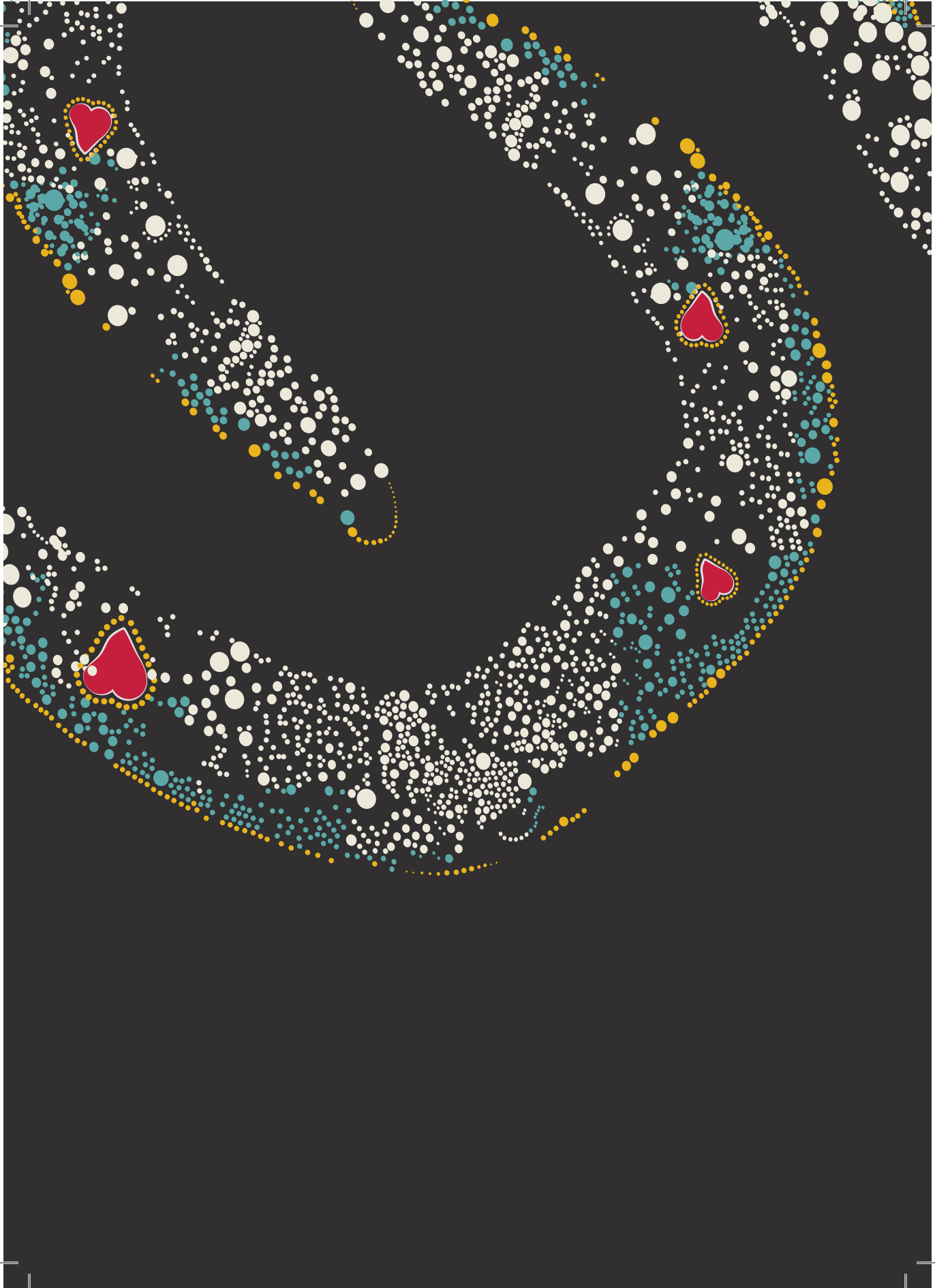
Y67A10A.1	oac-56	5,07E-10	5,14
R151.5	dpy-31	5,74E-10	1,80
Y68A4B.1	clec-243	5,80E-10	4,76
T14B4.5		6,80E-10	2,97
R04D3.1	cyp-14A4	7,11E-10	3,24
C35A5.2	ugt-33	8,49E-10	4,19
F59A7.2		1,15E-09	2,37
C06A6.5		1,65E-09	2,31
T03F7.2	srh-45	2,11E-09	5,76
B0238.13		2,30E-09	3,37
R57.2		2,65E-09	3,56
F25D1.3		3,03E-09	4,53
F15B9.2	far-4	3,05E-09	6,33
H14E04.1		3,99E-09	3,35
ZK596.1		4,22E-09	1,85
F39E9.2	btb-16	4,83E-09	2,18
C33C12.3	gba-1	4,98E-09	2,39
F19H8.4	mltn-9	5,04E-09	3,61
C16C4.15	math-10	5,64E-09	4,83
C18D11.6		6,32E-09	7,05
C44B7.7		6,79E-09	1,62
Y57E12B.3	lipl-6	7,99E-09	3,46
F27D9.6	dhs-29	1,03E-08	2,03
Y110A2AL.9		1,27E-08	2,82
F11E6.4		1,68E-08	3,23
W04C9.3	cutl-13	1,91E-08	3,91
F44G3.2		2,10E-08	3,61
F59B10.5		2,21E-08	2,78
F59B10.1	pqn-47	2,21E-08	2,80
R09E10.7	pqn-55	2,28E-08	1,91
K12G11.3	sodh-1	2,41E-08	2,82
T17A3.7	fbxb-84	2,41E-08	6,83
C03G6.14	cyp-35A1	2,46E-08	3,87
F11E6.9		2,66E-08	3,76
Y39A3B.7		2,66E-08	7,09
F54D8.1	dpy-17	2,67E-08	3,68
AC3.5		3,23E-08	2,21
M88.6	pan-1	3,54E-08	2,22
C33A12.6	ugt-21	3,60E-08	2,55
C04F6.5	dhs-27	3,72E-08	2,51
C05E4.14	srh-2	3,75E-08	3,32
Y22D7AL.9		4,35E-08	2,25
K01A2.2	far-7	4,50E-08	1,96

Table S2. Top 100 genes downregulated in polyglutamine-expressing animals at the L4 stage (blue data points from Figure 1F).

Sequence Name	Gene	False discovery rate (FDR)	Fold change
F22F4.4		2,89E-87	-10,36
T09F5.9	clec-47	3,55E-73	-6,09
C30G12.2		6,85E-33	-4,66
F52F10.4	oac-32	1,03E-25	-4,18
F45D11.15		1,41E-21	-6,41
F45D11.14		2,76E-20	-5,64
F22F4.1		7,22E-20	-7,00
F45D11.16		2,32E-19	-5,78
C06G8.1	swt-3	9,43E-17	-2,57
R17.3		5,28E-14	-2,57
K02E2.8		3,62E-13	-4,82
T04C9.4	mlp-1	1,37E-12	-1,70
C04G6.2		2,29E-12	-5,43
Y26D4A.2	hpo-2	3,21E-12	-5,66
ZC581.3		3,99E-12	-2,12
T19A6.4		7,75E-12	-2,68
C41H7.7	clec-3	7,82E-12	-2,96
C18H9.5		7,95E-12	-4,05
T21E8.1	pgp-6	8,63E-12	-2,60
E01G4.3		8,70E-12	-1,70
H32K16.2		4,02E-11	-3,53
T02B11.3		5,28E-11	-1,54
C31C9.7		6,31E-11	-2,07
C35D10.14	clec-5	1,13E-10	-2,55
F23F1.6		1,35E-10	-3,15
W09G10.6	clec-125	1,89E-10	-3,78
R11D1.3		2,05E-10	-3,31
W10G11.3		2,23E-10	-4,74
F07C6.3		2,87E-10	-2,60
K02E11.7		3,88E-10	-2,95
F17C11.5	clec-221	3,88E-10	-3,27
H06H21.8		1,50E-09	-1,67
C29F3.5	clec-230	1,56E-09	-4,35
T07F10.1		7,22E-09	-1,97
C05D9.3		8,65E-09	-1,38
K01C8.3	tdc-1	2,12E-08	-1,65
C10H11.6	ugt-26	2,66E-08	-1,89
K02E2.2	grd-11	3,32E-08	-3,75
C23G10.6		5,55E-08	-2,23
D1025.2	gcsH-1	5,87E-08	-2,07
C12D5.9		7,17E-08	-2,42
Y43F8C.1	nlp-25	8,13E-08	-3,41
K09E2.3		8,69E-08	-1,24
C05E7.2		1,55E-07	-3,01
C28A5.3	nex-3	3,46E-07	-1,67
T14G12.3	tag-18	7,57E-07	-1,18
F14H8.1	obr-2	7,86E-07	-1,29
F56F3.2	ndg-4	8,43E-07	-1,25
C41G11.1		9,56E-07	-2,32
T13F3.6		9,62E-07	-3,11
K11G9.3		1,10E-06	-2,55
R07E3.4		1,10E-06	-1,44
F02D8.4		1,20E-06	-2,07
Y27F2A.3	sri-40	1,21E-06	-2,97
C04C11.1		1,50E-06	-2,79
F23A7.8		1,60E-06	-1,44
T22B7.3		1,70E-06	-2,74
AC7.1	tkr-3	1,97E-06	-2,44
H02K04.1	clec-229	2,98E-06	-4,05
ZK550.2		3,07E-06	-2,43

Transcriptional profiling in *C. elegans* reveals developmental and cellular responses
to polyglutamine expansion proteins

Y39G10AR.6	ugt-31	3,13E-06	-1,91
T10B9.1	cyp-13A4	3,39E-06	-2,18
K05B2.3	ifa-4	3,45E-06	-1,47
F54B11.11		4,12E-06	-2,89
T09B9.1		5,15E-06	-2,64
Y38E10A.5	clec-4	5,37E-06	-2,29
C05E11.4	amt-1	5,58E-06	-3,77
T21E8.2	pgp-7	5,80E-06	-2,11
C05A9.1	pgp-5	7,05E-06	-2,02
F43C11.3	decr-1.1	7,05E-06	-4,01
C15F1.1		7,21E-06	-1,96
C01G6.9		8,34E-06	-2,34
Y67D8C.8	cpg-9	1,17E-05	-1,19
ZC266.1		1,25E-05	-2,20
R04B5.9	ugt-47	1,33E-05	-1,09
R09H10.7		1,65E-05	-1,80
C42D4.3		1,73E-05	-2,60
R09F10.7	pqn-57	1,82E-05	-3,15
T05B11.4		1,83E-05	-1,98
F09G8.5		1,92E-05	-1,71
F56D1.6	cex-1	1,96E-05	-1,86
T18H9.1	grd-6	2,07E-05	-2,63
AC3.4	pqn-2	2,22E-05	-2,70
C36B1.1	cle-1	2,96E-05	-0,98
D2096.4	sqv-1	3,18E-05	-1,11
T22B11.1		3,22E-05	-1,86
F07C3.7	aat-2	3,61E-05	-1,25
F38E11.1	hsp-12.3	3,73E-05	-1,90
T06E4.8		3,99E-05	-2,78
K07A1.13		4,07E-05	-1,62
T15B7.1		4,24E-05	-1,15
R05F9.5	gst-9	4,29E-05	-1,54
AC3.3	abu-1	4,35E-05	-2,66
ZC416.6		4,91E-05	-1,36
F53G12.5	mex-3	5,27E-05	-1,31
C11E4.1	gpx-5	5,87E-05	-1,50
C42C1.7	oac-59	6,35E-05	-1,70
H02I12.6	his-66	6,35E-05	-1,73
F15H10.2	col-13	6,46E-05	-2,38
C03A7.7	abu-6	7,17E-05	-2,35





CHAPTER VI

Loss of *Serf2* in mice results in embryonic lethality with incomplete penetrance

*Olga Sin^{1,2}, Esther Stroo¹, Wytse Hogewerf¹, Bart van de Sluis³,
Jan M. van Deursen⁴, Ellen A. A. Nollen¹*

¹ European Research Institute for the Biology of Aging, University of Groningen,
University Medical Centre Groningen, Groningen, The Netherlands

² Graduate Program in Areas of Basic and Applied Biology, Abel Salazar Biomedical
Sciences Institute, University of Porto, Porto, Portugal

³ Department of Paediatrics, Molecular Genetics Section, University of Groningen,
University Medical Centre Groningen, Groningen, The Netherlands.

⁴ Department of Paediatric and Adolescent Medicine, Mayo Clinic, Rochester,
Minnesota, USA

Work in progress

Abstract

Neurodegenerative disorders such as Parkinson's, Alzheimer's and polyglutamine diseases are commonly characterized by the presence of protein aggregates in specific neurons. How the formation of these proteinaceous inclusions occurs is still not fully understood, nor is their role in pathogenesis. A positive regulator of aggregate formation has been discovered in *C. elegans* models for aggregation of alpha-synuclein, amyloid-beta and polyglutamine. This protein, known as modifier of aggregation MOAG-4, has two human orthologs, SERF1A and SERF2, which drive polyglutamine aggregation in human cells. This suggests that SERF may drive disease-associated protein aggregation in the mammalian brain. To test this hypothesis, we generated a *Serf2* knockout mouse model in a C57BL/6 background. We found that the homozygous null allele is not inherited in Mendelian proportion, but some individuals are able to survive until adulthood. Our results suggest that full body knockout of *Serf2* is embryonic lethal with incomplete penetrance. Since such mice will be of limited use to study the function of *Serf2* in adult mice, we generated mice harboring a brain-specific knockout for *Serf2*. We report that the ablation of *Serf2* expression from the brain resulted in viable and fertile animals.

We conclude that *Serf2* is required for normal mouse development. Mice harboring the *Serf2* conditional knockout allele can now be used to study the biological function of *Serf2* in mouse development. Additionally, by mating these brain-specific knockout mice with mouse models of neurodegenerative diseases, we can also investigate whether *Serf2* is indeed a genetic modifier of protein aggregation.

Introduction

While protein aggregation and toxicity are hallmarks of neurodegenerative diseases such as Parkinson's (PD), Alzheimer's (AD) and polyglutamine diseases, their role in these diseases is currently unknown (reviewed in [1, 2]). To better understand the role of protein aggregation in disease, efforts have been made to identify genes and pathways that regulate protein aggregation. The identification of such genes has been aided by genetic screens in *Caenorhabditis elegans* models expressing aggregation-prone proteins, including alpha-synuclein, amyloid-beta and polyglutamine [3-10]. The function of the genetic modifiers identified in these screens has been shown to be conserved in human cell models and in mice [5, 11, 12].

In *C. elegans* models for PD, AD and polyglutamine diseases, we previously identified MOAG-4/SERF as a regulator of age-related proteotoxicity [5]. We showed that MOAG-4 promotes polyglutamine aggregation and toxicity and that its function is evolutionarily conserved in the mammalian orthologues SERF1A and SERF2 (small EDRK rich factor 1A and 2) [5]. In human cell models expressing polyglutamine, overexpression of MOAG-4/SERF enhanced aggregation and cell death [5].

The molecular function of the SERF proteins is unknown. These proteins are ubiquitously expressed and are therefore predicted to have a role in general cellular pathways. SERF1A was first identified in a comparative genomics study, where it was found to be a genetic modifier of spinal muscular atrophy in human patients [13]. A transcriptome analysis later revealed that SERF1A is downregulated in the cerebellum of HD patients [14]. Recently, SERF1A was shown to promote amyloid aggregation *in vitro* [15].

Here we aimed to test the hypothesis that SERF2 drives disease-associated protein aggregation in the mammalian brain. To this end, we generated *Serf2* homozygous null mutant mice. By intercrossing *Serf2*^{+/-} mice, we discovered that the segregation of the knockout allele did not follow Mendelian inheritance, suggesting that the targeted disruption of *Serf2* has fundamental implications for mouse development. To overcome this limitation, we generated brain-

specific *Serf2* knockout mice that are viable and fertile. We further investigated the distribution of *Serf2* expression and found that *Serf2* is transcribed in various tissues, namely brain, heart, intestine, kidney and liver.

Results and Discussion

Deletion of *Serf2* results in embryonic lethality with incomplete penetrance

To study the function of *Serf2*, we generated a mouse strain with an integrated promoter-driven selection cassette (Figure 1A) [16]. *Serf2* has 3 exons and spans about 4 kb in chromosome 2. Two *loxP* sites were inserted around exon 2 of the *Serf2* gene and its expression was prevented by the presence of an intronic *Neo* selection cassette that was surrounded by two *loxP* sites and two *FRT* sites (Figure 1A). *Hprt* Cre-induced recombination excised the *Neo* gene and exon 2 of *Serf2*, thereby generating a heterozygous reporter knockout for *Serf2*, with the *lacZ* reporter transgene under the control of the *Serf2* promoter (*Serf2*^{+/-}, Figure 1A, B).

We next attempted to generate *Serf2* null mice (*Serf2*^{-/-}) by intercrossing *Serf2*^{+/-} mice. The numbers of *Serf2*^{-/-} progeny obtained were far lower than those

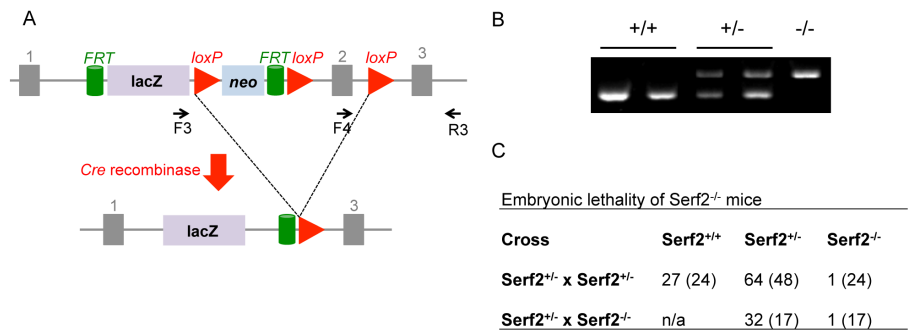


Figure 1. Gene targeting strategy for the *Serf2* knockout model. (A) Diagram of the *Serf2* targeting vector. The *Serf2* exon 2 is flanked by two *loxP* sites. Expression is prevented by an intronic *Neo* selection cassette surrounded by two *loxP* sites and two *FRT* sites. Mating with a *Hprt* Cre deleter mouse strain removes exon 2, resulting in a reporter knockout whereby the *lacZ* transgene is expressed under the control of the *Serf2* promoter. Diagram adapted from [16]. (B) Genotype of wild type, heterozygous and knockout *Serf2* transgenic mice. Primers F3 and R3 were used to distinguish the targeted allele (420 bp) from the untargeted wild type allele (335 bp), identified with primers F4 and R3. (C) Numbers of mice obtained for each genotype from intercrossing *Serf2*^{+/-} mice or by mating *Serf2*^{+/-} mice with *Serf2*^{-/-} mice. Numbers in parentheses represent the expected Mendelian ratio for each genotype.

expected according to Mendelian inheritance and only a small proportion was able to survive throughout adulthood, suggesting embryonic lethality with incomplete penetrance (Figure 1B, C). When we attempted to mate *Serf2*^{+/-} mice with *Serf2*^{-/-} mice, *Serf2*^{-/-} offspring of this cross were also underrepresented (Figure 1C). Together, these results demonstrate that *Serf2* is essential for mouse development.

Serf2 is expressed in disease-related brain areas

To understand the role of *Serf2*, we next wanted to identify the tissues in which *Serf2* is normally expressed in *Serf2*^{+/+} and *Serf2*^{+/-} mice. To do so, primers specific for *Serf2* were used to investigate its expression in the brain, heart, intestine, kidney, liver and pancreas by qPCR (Figure 2A). *Serf2* was expressed in most the tissues examined in this study. mRNA expression in *Serf2*^{+/-} mice

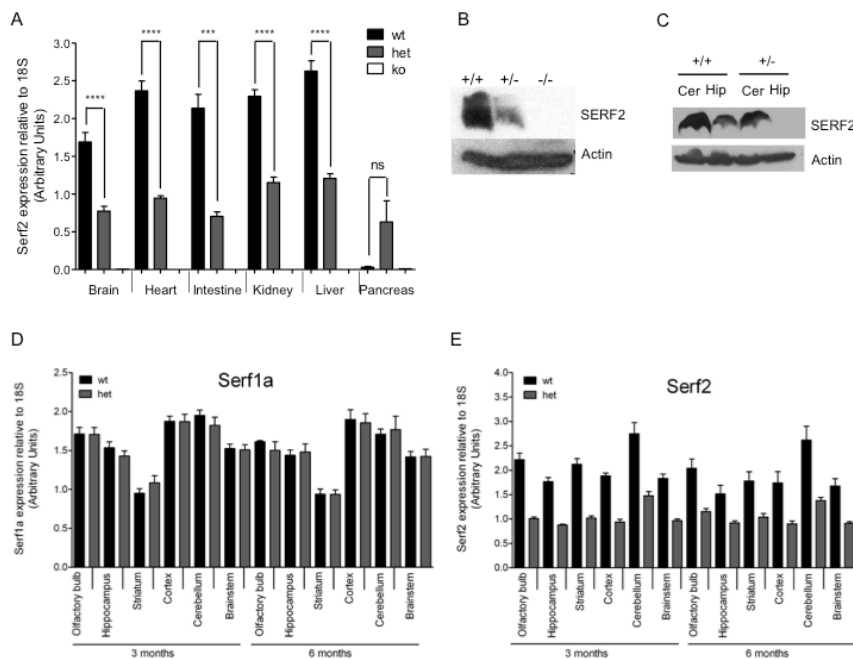


Figure 2. Expression of *Serf2* in different tissues. (A) Expression of *Serf2* in tissues of 3-month-old mice. Relative mRNA levels were measured by qPCR in wild type (n=3), heterozygous (n=8) and knockout mice (n=1) for *Serf2*. **(B)** Protein expression of *Serf2* in liver samples obtained from wild type, heterozygous and knockout mice by Western blot. **(C)** Protein expression of *Serf2* in the cerebellum (Cer) and hippocampus (Hip) of wild type and heterozygous mice by Western blot. In both experiments, actin was used as a loading control. **(D)** Expression of *Serf1a* in different brain regions of wild type mice (n=5) and mice heterozygous for the *Serf2* knockout allele (n=5). **(E)** Expression of *Serf2* in different brain regions of wild type mice (n=5) and mice heterozygous for the *Serf2* knockout allele (n=5). Relative mRNA levels were measured in the olfactory bulb, hippocampus, striatum, cortex, cerebellum and brainstem. **** p<0.0001; *** p<0.001; ns is not significant.

was significantly lower (up to 50%) than that in control mice (Figure 2A). This was later confirmed by measuring *Serf2* protein levels in the brain and liver (Figure 2B, C).

Knowing that *Serf2* was expressed in the brain and aiming to use this model to later generate a *Serf2* knockout in a neurodegenerative disease background, we asked in which brain areas *Serf2* was expressed. We found that *Serf2* was expressed in different brain regions, including the brainstem, the olfactory bulb, the cerebellum, the striatum and, importantly, the hippocampus and the cortex, which are frequently affected in neurodegenerative diseases (Figure 2E). We examined 3 and 6-month old mice and confirmed that *Serf2*^{+/-} mice had reduced expression of *Serf2* while *Serf1a* expression was unaltered (Figure 2D, E). A reduction in expression of *Serf2* was not compensated by the expression of *Serf1a*, supporting the use of this model to study *Serf2* function. Our findings show that *Serf2* is expressed in different tissues, including the brain. Specifically, *Serf2* is expressed in brain areas relevant to disease and loss of *Serf2* does not affect *Serf1a* expression.

Mice with brain-specific knockout for *Serf2* are viable and fertile

Since the full body knockout of *Serf2* resulted in embryonic lethality, we could not take advantage of these mice to study the role of *Serf2* in protein aggregation in the brain. To overcome this limitation, we specifically eliminated *Serf2* expression in the brain. First, mice harboring the *Serf2* construct were crossed with *FLP* deleter mice to generate a conditional *Serf2* allele (*Serf2*^{flox/flox}, Figure 3A). *Serf2*^{flox/flox} mice express the transgene as wild type and have normal appearance (data not shown). To specifically eliminate *Serf2* expression in the brain, *Serf2*^{flox/flox} mice were mated with mice that express the *Cre* recombinase under the control of the *Sox1* promoter, which restricts expression of *Cre* to the central nervous system (Figure 3A). Conditional knockout mice (*Cre Sox1*⁺ *Serf2*^{-/-}) derived from this cross were viable and fertile (Esther Stroo, personal communication). To demonstrate that *Serf2* expression was eliminated only in the brain, we measured expression levels of *Serf2* in the brain and compared them with those in other tissues, including the kidney and the liver. When compared with expression in *Serf2*^{flox/flox} control mice, expression of *Serf2* in *Cre Sox1*⁺ *Serf2*^{-/-} mice was almost completely abolished

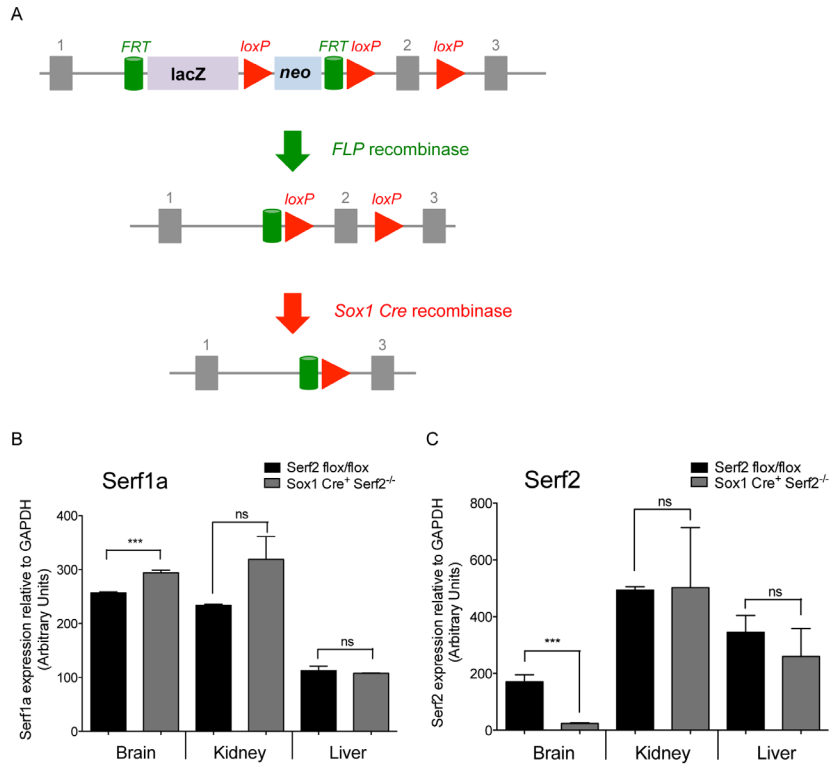


Figure 3. Gene targeting strategy for the brain-specific knockout of *Serf2*. (A) To generate a conditional allele for *Serf2*, mice harboring the *Serf2* targeting vector were mated with a *FLP* deleter line to excise the *lacZ* transgene and the *Neo* selection cassette. This resulted in a floxed *Serf2* allele (*Serf2*^{flox/flox}), in which *Serf2* is expressed as the wild type. Next, *Serf2*^{flox/flox} mice were mated with *Sox1 Cre* deleter mice to generate *Cre Sox1*⁺ *Serf2*^{-/-} progeny. Diagram adapted from [16]. (B) Quantitative PCR of *Serf1a* in the brain (n=4), the liver (n=2) and the kidney (n=2) of *Serf2*^{flox/flox} and *Cre Sox1*⁺ *Serf2*^{-/-} mice (5 months old). (C) Quantitative PCR of *Serf2* in the brain (n=4), the liver (n=2) and the kidney (n=2) of *Serf2*^{flox/flox} and *Cre Sox1*⁺ *Serf2*^{-/-} mice (5 months old). *** p<0.001; ns is not significant.

in the brain but unaltered in the kidney and liver (Figure 3C). Together, these results demonstrate the successful knockout of *Serf2* in the brain without its expression in other tissues being affected.

To investigate whether the knockout of *Serf2* affected *Serf1a* expression, we also measured *Serf1a* in the brain. We detected a statistically significant increase of *Serf1a* in the brain of *Cre Sox1*⁺ *Serf2*^{-/-} mice, which suggests a compensation mechanism by *Serf1a* in the absence of *Serf2* (Figure 3B).

Conclusion

In summary, *Serf2* is essential for mouse development. We describe the generation of brain-specific *Serf2* knockout mice that can be used not only to further study the role of *Serf2* in mouse development but also to determine *Serf2* involvement in proteotoxicity by mating these knockout mice with mouse models of neurodegenerative diseases.

Methods

Expression construct and generation of transgenic mice

The *Serf2* targeting vector was generated by the International Knockout Mouse Consortium (IKMC, Figure 1A) [16, 17]. Briefly, the exon 2 of *Serf2* is flanked by two *loxP* sites its expression is prevented by an intronic *Neo* selection cassette, which is flanked by two *loxP* sites and two *FRT* sites. The construct was injected into C57BL/6N embryonic stem cells [18]. Two founder lines (B11 and G9) were positive for the construct and subsequently expanded for this study. To generate a full body knockout of *Serf2*, both lines were crossed with a *Hprt Cre* deleter mouse strain to remove the *PGK-Neo* cassette and *Serf2* floxed exon. Confirmation of correct recombination was performed by PCR genotyping analysis of ear biopsy. All mice were heterozygous with respect to the construct. Animals were backcrossed into the C57Bl/6J background (Charles River) six times.

Genotyping

DNA was prepared from ear biopsy and processed with *prepGEM* Tissue (#PTI0500, ZyGEM Corporation Ltd). PCR reactions contained three primers, one sense primer specific for the *Serf2* transgene F3 (5'-CCGGTCGCTACCATTACCAG-3'); a second sense primer specific for genomic *Serf2* F4 (5'-GATGATGGGCTTTCTGCTGC-3'); and one antisense primer present in the transgene and mouse *Serf2* R3 (5'-CTTGATATGTGAAGCCCCTGC-3'). The knockout allele was identified with pair F3 and R3, generating a 420 bp product; and the wild type allele was identified with pair F4 and R3, generating a 335 bp product. Cycling conditions were 2 min at 94°C; 35 cycles of 30 sec at 94°C; 30 sec at 60°C; 1 min at 72°C; and 7 min at 72°C.

Quantitative PCR

Total RNA was extracted using TRIzol Reagent (Life Technologies) according to the manufacturer's description. Total RNA quality and concentration were assessed using a NanoDrop 2000 spectrophotometer (Thermo Scientific/Isogen Life Science). cDNA was made from 1.5 ug (tissues) or 2 ug (brain regions) total RNA with a RevertAid H Minus First Strand cDNA Synthesis kit (Life Technologies) using random hexamer primers.

Quantitative real-time PCR was performed using a Roche LightCycler 480 Instrument II (Roche Diagnostics) with SYBR green dye (Bio-Rad Laboratories) to detect DNA amplification. Relative transcript levels were quantitated using a standard curve of pooled cDNA solutions. Expression levels were normalized to 18S mRNA levels. The primers for RT-PCR used were Serf1 F2 (5'-TGGCCCGTGGAATCAAAGAGAAA-3'); Serf1 R2 (5'-TGCATGATCTCTGAATCCCTCTGCT-3'); Serf2 F2 (5'-CCGCGGTAACCAGCGAGAGC-3'); Serf2 R2 (5'-TCCGAGTCCCTCTGCTTGCG-3'); 18S F1 (5'-CGGACAGGATTGACAGATTG-3'); 18S R1 (5'-CAAATCGCTCCACCACTAA-3').

Western Blot

Frozen tissue was homogenized in RIPA buffer (10mM Tris-Cl pH 8.0; 1mM EDTA; 0.5mM EGTA; 1% Triton X-100; 0.1% sodium deoxycholate; 0.1% SDS; 140mM NaCl). Samples were homogenized with a tissue grinder pestle and incubated on ice for 30 min, followed by centrifugation at high speed for 30 min at 4°C. The supernatant was transferred to a new tube and protein was quantified using the Pierce BCA Protein Assay kit (#23225, Life Technologies). Approximately 130 ug (cerebellum and hippocampus) or 40 ug (liver) of protein were loaded onto 15% SDS-PAGE gels. Proteins were transferred to nitrocellulose membranes and blocked with 5% BSA in PBS-T. Membranes were incubated with primary antibodies for SERF2 at 1:1500 dilution overnight (#11691-1-AP, Proteintech) or actin at 1:10,000 dilution overnight (#3134S, Cell Signaling Technology). Washes were performed with PBS-T. Incubation with secondary anti-rabbit for SERF2 or anti-mouse for actin was done at 1:10,000 dilution for 1 hour at room temperature. Antibody binding was visualized with an ECL kit (Amersham).

Acknowledgements

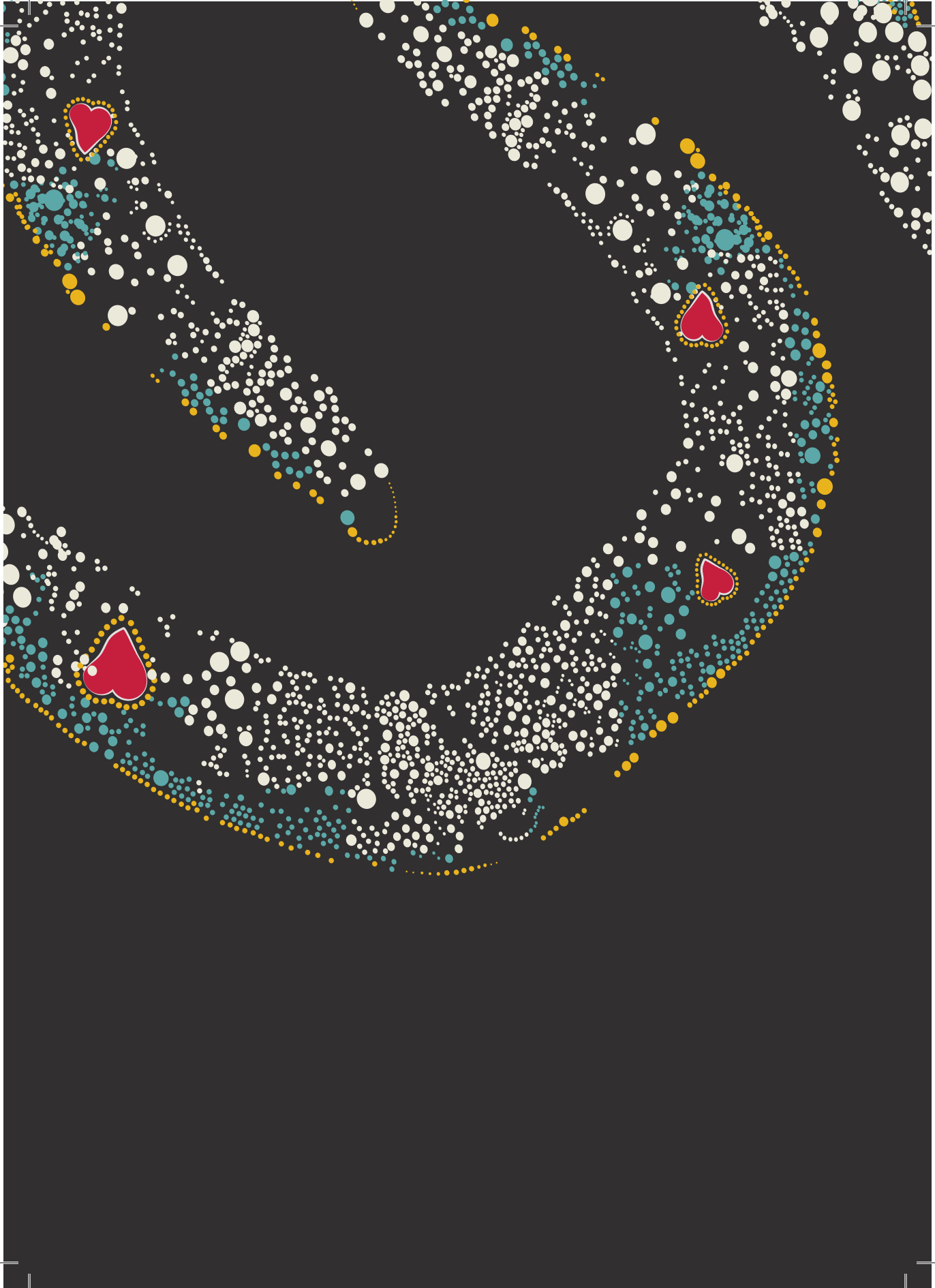
We thank Rui Costa and Catherine French for the *Sox1 Cre* transgenic mice. We thank Mathias Jucker and Ehud Cohen for fruitful discussion. We thank Sally Hill for editing the manuscript. Alumni chapter Gooische Groningers facilitated by the Ubbo Emmius Fund. This work was supported by the European Research Council (ERC) starting grant (to E.A.A.N.), the Ubbo Emmius Fonds and the Fundação para a Ciência e Tecnologia (SFRH/BPD/51009/2010) (to O. S.).

Author contributions

O. S., W. H. and E. A. A. N. designed the experiments. O. S. and W. H. performed the experiments. O. S. and E. A. A. N. wrote the manuscript.

References

1. Knowles TPJ, Vendruscolo M, Dobson CM (2014) The amyloid state and its association with protein misfolding diseases. *Nature Publishing Group* 15:384–396. doi: 10.1038/nrm3810
2. Soto C (2003) Unfolding the role of protein misfolding in neurodegenerative diseases. *Nat Rev Neurosci* 4:49–60. doi: 10.1038/nrn1007
3. van der Goot AT, Zhu W, Vazquez-Manrique RP, et al. (2012) Delaying aging and the aging-associated decline in protein homeostasis by inhibition of tryptophan degradation. *Proc Natl Acad Sci USA* 109:14912–14917. doi: 10.1073/pnas.1203083109
4. Silva MC, Fox S, Beam M, et al. (2011) A genetic screening strategy identifies novel regulators of the proteostasis network. *PLoS Genet* 7:e1002438. doi: 10.1371/journal.pgen.1002438
5. van Ham TJ, Holmberg MA, van der Goot AT, et al. (2010) Identification of MOAG-4/SERF as a regulator of age-related proteotoxicity. *Cell* 142:601–612. doi: 10.1016/j.cell.2010.07.020
6. Hamamichi S, Rivas RN, Knight AL, et al. (2008) Hypothesis-based RNAi screening identifies neuroprotective genes in a Parkinson's disease model. *Proc Natl Acad Sci USA* 105:728–733. doi: 10.1073/pnas.0711018105
7. Kuwahara T, Koyama A, Koyama S, et al. (2008) A systematic RNAi screen reveals involvement of endocytic pathway in neuronal dysfunction in alpha-synuclein transgenic *C. elegans*. *Human Molecular Genetics* 17:2997–3009. doi: 10.1093/hmg/ddn198
8. van Ham TJ, Thijssen KL, Breitling R, et al. (2008) *C. elegans* model identifies genetic modifiers of alpha-synuclein inclusion formation during aging. *PLoS Genet* 4:e1000027. doi: 10.1371/journal.pgen.1000027
9. Kraemer BC, Burgess JK, Chen JH, et al. (2006) Molecular pathways that influence human tau-induced pathology in *Caenorhabditis elegans*. *Human Molecular Genetics* 15:1483–1496. doi: 10.1093/hmg/ddl067
10. Nollen EAA, Garcia SM, van Haaften G, et al. (2004) Genome-wide RNA interference screen identifies previously undescribed regulators of polyglutamine aggregation. *Proceedings of the National Academy of Sciences* 101:6403–6408. doi: 10.1073/pnas.0307697101
11. Komatsu M, Wang QJ, Holstein GR, et al. (2007) Essential role for autophagy protein Atg7 in the maintenance of axonal homeostasis and the prevention of axonal degeneration. *Proceedings of the National Academy of Sciences* 104:14489–14494. doi: 10.1073/pnas.0701311104
12. Teuling E, Bourgonje A, Veenje S, et al. (2011) Modifiers of mutant huntingtin aggregation: functional conservation of *C. elegans*-modifiers of polyglutamine aggregation. *PLoS Curr* 3:RRN1255. doi: 10.1371/currents.RRN1255
13. Scharf JM, Endrizzi MG, Wetter A, et al. (1998) Identification of a candidate modifying gene for spinal muscular atrophy by comparative genomics. *Nat Genet* 20:83–86. doi: 10.1038/1753
14. Hodges A (2006) Regional and cellular gene expression changes in human Huntington's disease brain. *Human Molecular Genetics* 15:965–977. doi: 10.1093/hmg/ddl013
15. Falsone SF, Meyer NH, Schrank E, et al. (2012) SERF protein is a direct modifier of amyloid fiber assembly. *Cell Rep* 2:358–371. doi: 10.1016/j.celrep.2012.06.012
16. Skarnes WC, Rosen B, West AP, et al. (2011) A conditional knockout resource for the genome-wide study of mouse gene function. *Nature* 474:337–342. doi: 10.1038/nature10163
17. Osterwalder M, Galli A, Rosen B, et al. (2010) Dual RMCE for efficient re-engineering of mouse mutant alleles. *Nature Methods* 7:893–895. doi: 10.1038/nmeth.1521
18. Pettitt SJ, Liang Q, Rairdan XY, et al. (2009) Agouti C57BL/6N embryonic stem cells for mouse genetic resources. *Nature Methods* 6:493–495. doi: 10.1038/nmeth.1342





CHAPTER VII

Discussion and perspectives



Many neurodegenerative diseases are characterized by clinicopathological features that affect motor capacity and cognition in affected patients. Examples of such diseases are Alzheimer's, Parkinson's, polyglutamine diseases and amyotrophic lateral sclerosis (ALS). Unfortunately, no cure is currently available and the risk of developing these diseases increases with aging. Interestingly, a hallmark of post-mortem biopsies from the brains of affected patients is the presence of protein aggregates in neurons. While specific aggregation-prone proteins have been identified for each neurodegenerative disease, their contribution to disease remains unresolved. More specifically, we do not know what triggers protein aggregation and how this is mechanistically linked to pathogenesis. What we do know is that protein homeostasis is affected in these diseases, and understanding how it is altered may help us not only to identify the cellular processes involved but also – in the long term – to target them for potential disease-modifying treatments.

The aim of the work presented in this thesis was to dissect the cellular and molecular aspects of protein aggregation, in order to find mechanisms that help to explain how protein homeostasis correlates with pathogenesis. A main goal was to discover genetic regulators of protein aggregation in an unbiased and hypothesis-free manner, with the expectation of identifying new regulators and therefore new mechanisms in protein aggregation.

We were successful in identifying one such new regulator in the form of MOAG-2/LIR-3. **Chapter IV** describes the identification and characterization of this modifier of aggregation in a *C. elegans* model of polyglutamine diseases. We discovered that MOAG-2/LIR-3 is an RNA Polymerase (Pol) III-associated transcriptional regulator that drives polyglutamine aggregation. In this chapter we propose a mechanism whereby aggregation-prone proteins recruit normally functioning cellular proteins to promote their own aggregation. Despite this important finding, this study leaves several questions unanswered – these are discussed below, together with suggestions for directions of future research in the context of MOAG-2/LIR-3 and the link between the non-coding genome and protein aggregation.

1. Aggregation-prone proteins affect the non-coding genome

One surprising finding from this work was that in our model the presence of polyglutamine expansion proteins induced the downregulation of small non-coding RNAs (ncRNAs) – specifically snRNAs, snoRNAs and tRNAs – demonstrating that the non-coding genome is also affected by the presence of aggregation-prone proteins. It is still unknown how this downregulation occurs and we propose two different possibilities that could explain this observation. The first possibility is that misfolded, aggregation-prone proteins directly sequester these small ncRNAs into aggregates, thereby reducing their abundance in the cell. The second possibility is that aggregation-prone proteins sequester MOAG-2/LIR-3, subsequently inhibiting it from driving small ncRNA transcription. To experimentally address these hypotheses, it would be necessary to isolate the polyglutamine aggregates and detect the presence of small ncRNAs (e.g. qPCR) or MOAG-2/LIR-3 (e.g. co-immunoprecipitation).

The next unresolved question is what biological consequences the downregulation of small ncRNAs might have. For example, the reduction of one type of small ncRNAs in the nucleus – the U1 snRNAs – affects the splicing of premature mRNA [1, 2]. In our case, however, downregulation of snRNAs did not affect the splicing of premature mRNA (data not shown). Our observations so far point to the notion that it is possible to retain functional splicing events despite lower levels of splicing factors. The downregulation of another type of small ncRNA – the tRNAs – may also have biological consequences, and it will be interesting to investigate how tRNA function is affected, if at all (e.g. is translation efficiency maintained? does the proteome change?).

Finally, it will also be worth assessing whether the downregulation of small ncRNAs is cell type-specific (e.g. do polyglutamine-expressing cells have a distinct ncRNA repertoire?) and whether altered ncRNA homeostasis is specific to polyglutamine expansion proteins or shared by other aggregation-prone proteins (e.g. alpha-synuclein, amyloid-beta).

2. Is MOAG-2/LIR-3 related to calcium-regulated transcription factors?

Once we discovered that MOAG-2/LIR-3 was a transcription factor we were interested to know whether it shared homology with other mammalian transcription factors. Protein sequence comparisons revealed that MOAG-2/LIR-3 is predominantly conserved among the *Caenorhabditis* genus. It does, however, share 27% homology with the human calcineurin B homologous protein 1 (CHP1) and 37% homology with the human zinc finger protein 64 (ZNF64). The presence of these two functional motifs – a calcium-binding domain and a zinc finger – is typical of human calcium-regulated transcription factors, such as the cAMP-responsive element-binding protein (CREB)-binding protein (CBP), the downstream regulatory element antagonist modulator (DREAM) and FOXO3A [3]. The combination of these two domains – calcium binding and DNA binding – poses the interesting possibility that MOAG-2/LIR-3 may be a calcium-dependent transcription factor that regulates expression of small ncRNAs.

3. What are the proteins that interact with MOAG-2/LIR-3?

We can learn more about MOAG-2/LIR-3 function by identifying the types of proteins with which it interacts or forms complexes. We have shown that MOAG-2/LIR-3 binds to the same promoters as the RNA Pol III complex. We have also shown that inactivation of MOAG-2/LIR-3 reduces the levels of transcribed small ncRNAs, suggesting that MOAG-2/LIR-3 cooperates with the RNA Pol III complex to promote small ncRNA transcription. The question of whether MOAG-2/LIR-3 is physically part of the RNA Pol III complex could be answered by using purified MOAG-2/LIR-3 in co-immunoprecipitation experiments, followed by mass spectrometry. This approach would also enable us to identify other co-immunoprecipitated proteins and establish a protein interaction network for MOAG-2/LIR-3, which would tell us more about MOAG-2/LIR-3 function and whether it cooperates with other co-factors to regulate the expression of small ncRNA genes.

4. Could MOAG-2/LIR-3 be an RNA-binding protein?

The C2H2 zinc fingers of MOAG-2/LIR-3 are structurally homologous to the canonical C2H2 zinc fingers of the transcription factor IIIA (TFIIIA) in *Xenopus laevis* [4]. TFIIIA is a promoter-bound recruitment factor that recruits TFIIIB which, in turn, recruits Pol III to the promoters of genes encoding of 5S rRNA [5]. The binding of TFIIIA to the promoter takes place through physical binding of zinc fingers 4-9 to two specific regions: Box A and the intermediate element [6]. In addition to binding to 5S promoters, TFIIIA also binds to 5S RNA itself through zinc fingers 4-7 [7-9]. The fact that we found MOAG-2/LIR-3 to be frequently bound to 5S promoters suggests that it may also act as a recruitment factor, recruiting other components of the RNA Pol III complex to Box A and Box B. Furthermore, the resemblance of motifs within its structure to the zinc fingers of TFIIIA opens up the possibility that MOAG-2/LIR-3 might bind directly to rRNA or to any other type of RNA, a prospect that will have to be assessed in future studies.

5. Do aggregation-prone proteins alter binding of MOAG-2/LIR-3 to its transcriptional targets?

We have shown that MOAG-2/LIR-3 binds to the promoters of snRNA, snoRNA and tRNA genes to promote their transcription. One open question stemming from this work is how the presence of aggregation-prone proteins modifies the binding of MOAG-2/LIR-3 to its transcriptional targets. We propose that polyglutamine sequesters MOAG-2/LIR-3 or modifies its localization, thereby preventing MOAG-2/LIR-3 binding to its DNA targets such that it can no longer transcribe small ncRNAs with the same efficiency as that in wild type animals. This could explain why we observe lower levels of small ncRNAs in the presence of aggregation-prone proteins. Indeed, it would be interesting to verify whether MOAG-2/LIR-3 can still retain its function as a transcriptional regulator in polyglutamine-expressing animals, a hypothesis which will require testing in additional ChIP seq experiments.

The identification and characterization of MOAG-2/LIR-3 in *C. elegans* offers a major contribution to our understanding of the aggregation pathway in neurodegenerative diseases, and in particular how benign cellular proteins can convert their function to an aggregation-promoting factor. A second major

step focused on the effects of protein aggregation at the transcriptional level. In **Chapter V**, we showed that the presence of polyglutamine-expansion proteins induced the expression of stress response genes involved in: 1) oxidative stress; 2) the ER-associated unfolded protein response (UPR); and 3) the innate immune response. Oxidative stress results from the imbalance between the production and clearance of reactive oxygen species (ROS), which can cause damage to DNA, proteins and other cellular components (reviewed in [10]). Our results come in line with previous published work implicating oxidative stress in Huntington's disease [11, 12]. It would be interesting to further explore this correlation and determine what specific cellular events lead to oxidative stress (e.g. mitochondrial dysfunction? overproduction of ROS? insufficient antioxidant production?).

If unfolded proteins accumulate in the ER, the UPR is responsible for restoring protein homeostasis by fine-tuning protein-folding load with protein-folding capacity (reviewed in [13]). Previous published work showed that the UPR contributes to neuroprotection in a *C. elegans* model expressing alpha-synuclein in the dopaminergic neurons [14]. Although polyglutamine expansion proteins are not expected to enter the secretory pathway, it is still reasonable to consider that their accumulation at the ER (perhaps immediately after protein synthesis) would be sufficient to trigger the UPR as protective response to aggregation-prone proteins.

A surprising finding was the involvement of the innate immune response to polyglutamine expansion proteins. Several genes that encode antimicrobial peptides were upregulated in our study, suggesting that aggregation-prone proteins can also elicit a systemic effect in the organism. It is possible that aggregation-prone proteins are perceived by the organism as a pathogen and thereby elicit the immune system to counteract their toxicity. Indeed, the involvement of the cellular stress responses described here show that the organism takes advantage of distinct cellular defense mechanisms at its disposal to combat the pathogenic presence of aggregation-prone proteins. Following this line of thought, it would be interesting to investigate whether there is a cross talk between these cellular responses to aggregation-prone proteins.

Lastly, we saw that the effects of aggregation-prone proteins go beyond cellular stress responses and can also delay *C. elegans* development. Polyglutamine diseases are not known to be associated with impaired development, therefore suggesting that the developmental delay observed in our polyglutamine model may be specific to nematodes.

A final major step in our understanding of protein aggregation is well under way as a result of our work with MOAG-4 and its mammalian orthologs, SERF1A and SERF2. Below we discuss the contributions made by our study on SERF2 in the brains of mice (**Chapter VI**) and discuss plans for future experiments in mouse models of Alzheimer's disease.

The aggregation-prone protein most commonly associated with Alzheimer's disease (AD) is amyloid-beta (reviewed in [15, 16]). We and others have previously shown that aggregation of amyloid-beta is promoted both *in vitro* and *in vivo* by MOAG-4/SERF [17, 18]. In a *C. elegans* model of amyloid-beta, the transgene is expressed intracellularly in the body wall muscle, where it accumulates and contributes to progressive paralysis in the worm [19]. We have shown that deletion of MOAG-4 ameliorates this paralysis phenotype by reducing the amount of seeding-competent amyloid-beta [17]. In addition, our studies in human cell models have shown that the function of MOAG-4 is conserved in two mammalian orthologs, SERF1A and SERF2 [17]. This led us to hypothesize that SERF proteins could be driving amyloid-beta aggregation in the AD brain. To test this hypothesis, we generated a *Serf2* knockout mouse and crossed it with an AD mouse model. In **Chapter VI**, we discovered that a full body knockout of *Serf2* results in embryonic lethality with incomplete penetrance. Since this does not allow us to study the role of *Serf2* in adult mice, we generated brain-specific *Serf2* knockout mice instead. As a next step in this project, we plan to mate these animals with AD mouse models.

The aim of generating a brain-specific *Serf2* knockout mouse in an AD background is to find out whether the function of MOAG-4/SERF in proteotoxicity is conserved from *C. elegans* to mammals. Specifically, we want to know whether removing SERF from the brain in AD mouse models suppresses amyloid aggregation and neurotoxicity. We also want to find out whether

changes in proteotoxicity are accompanied by changes in motor and cognitive functions. The experimental design is depicted in Figure 1. Brain-specific *Serf2* knockout mice – *Sox1 Cre⁺ Serf2^{-/-}* – will be mated with two distinct AD mouse models. The first strategy involves mating the *Sox1 Cre⁺ Serf2^{-/-}* mice with the APPPS1 mouse model [20]. In this model, the human amyloid precursor protein (APP) harboring the K595N/M596L “Swedish” mutation (otherwise known as K670N/M671L) and the human presenelin 1 (PS1) protein harboring the L166P mutation are co-expressed under the Thy1 promoter, which directs transgene expression to the postnatal brain [20]. Both mutations are associated with early onset AD in humans. In this model, amyloid-beta deposition is seen as early as six weeks of age, with cognitive impairment appearing at 7 months of age [20, 21]. From the mating between the *Sox1 Cre⁺ Serf2^{-/-}* mice and the APPPS1 mice, we expect to obtain progeny in Mendelian proportions with the following characteristics: 1) wild type for all transgenes; 2) transgenic for APPPS1; 3) wild type lacking SERF2 in the brain; and 4) transgenic for APPPS1 and lacking SERF2 in the brain (Figure 1). Here, we will focus predominantly on brain histology and immunohistochemistry to study the effect of a lack of *Serf2* in the brain on amyloid-beta deposition (Figure 1).

In the second strategy, we will mate *Sox1 Cre⁺ Serf2^{-/-}* mice with the APPswe/PSEN1dE9 mouse model, from which we also expect to obtain all four desired phenotypes in Mendelian proportions (Figure 1) [22]. In this model of AD, a chimera of mouse and human APP harboring the K595N/M596L “Swedish” mutation (APPswe) is co-expressed with the exon-9-deleted (delta E9) variant of human mutated PS1 protein (PSEN1dE9), the latter also associated with early onset AD [22]. The expression of both transgenes is under the mouse prion promoter, which restricts their expression to the neurons of the central nervous system [22]. In APPswe/PSEN1dE9 mice, amyloid-beta plaques typically appear at 4-6 months of age, and plaque accumulation progresses up to 12 months of age, when cognitive deficits also start to appear [23, 24]. We will subject these mice to a battery of behavioral tests to assess spatial memory reference (e.g. Morris water maze), exploration and anxiety (e.g. open field) and motor capacity (e.g. grip strength) (Figure 1). In parallel, we will perform histological and immunohistochemical assays to assess amyloid deposition and neuronal toxicity. Since AD is a progressive disorder, we want to learn how aggregation

formation, neurotoxicity and cognition are affected during aging. For this reason, we will include both 3 to 4 and 6 to 7-month old animals in our analysis. In all, we hope to uncover the function of *Serf2* in the brain and determine whether it acts as a genetic modifier of proteotoxicity and cognition in AD. Results from these studies will provide insight into how the brain copes with protein aggregation during aging and whether this is accompanied by changes in behavior. Ultimately, this study may reveal *Serf2* to be a potential therapeutic target for modulation of neurodegeneration.

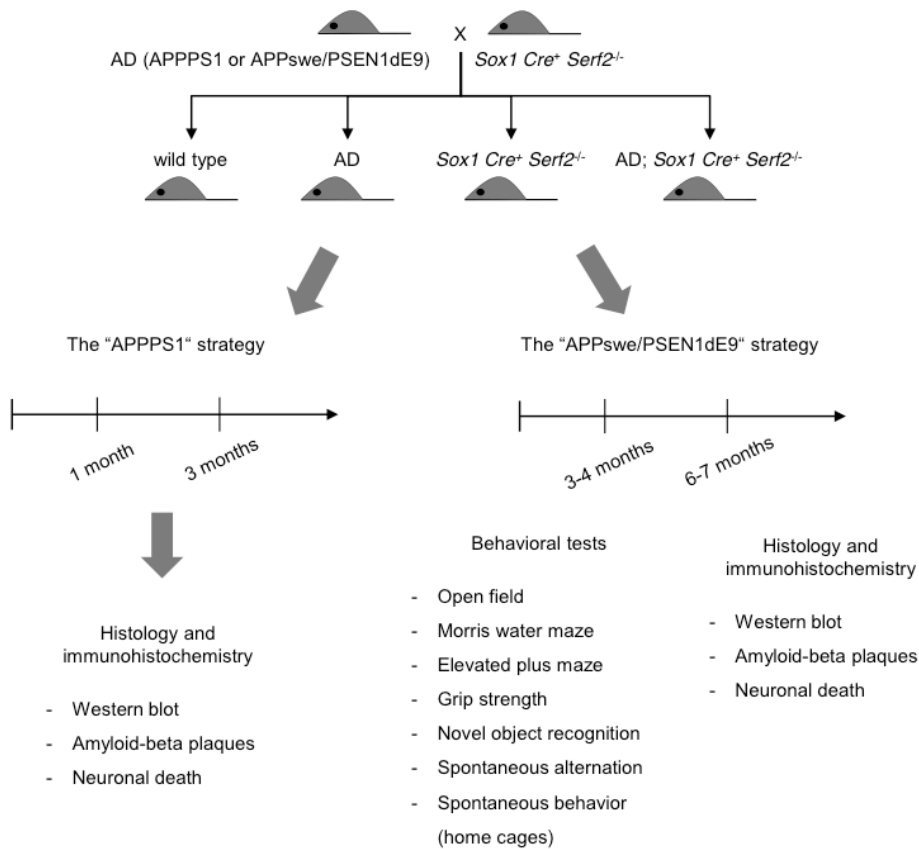


Figure 1. Diagram representing the generation of a brain-specific *Serf2* knockout mouse (*Sox1 Cre+ Serf2-/-*) in two different AD mouse models. In the "APPPS1" strategy, the brain-specific *Serf2* knockout mice will be mated with the APPPS1 mice and focus on histological and immunohistochemical analyses at both 1 and 3 months of age. In the "APPswe/PSEN1dE9" strategy, the brain-specific *Serf2* knockout mice will be mated with the APPswe/PSEN1dE9 mice. A set of behavioral tests will be performed parallel to histological and immunohistochemical analyses to 3 to 4 and 6 to 7-month old mice.

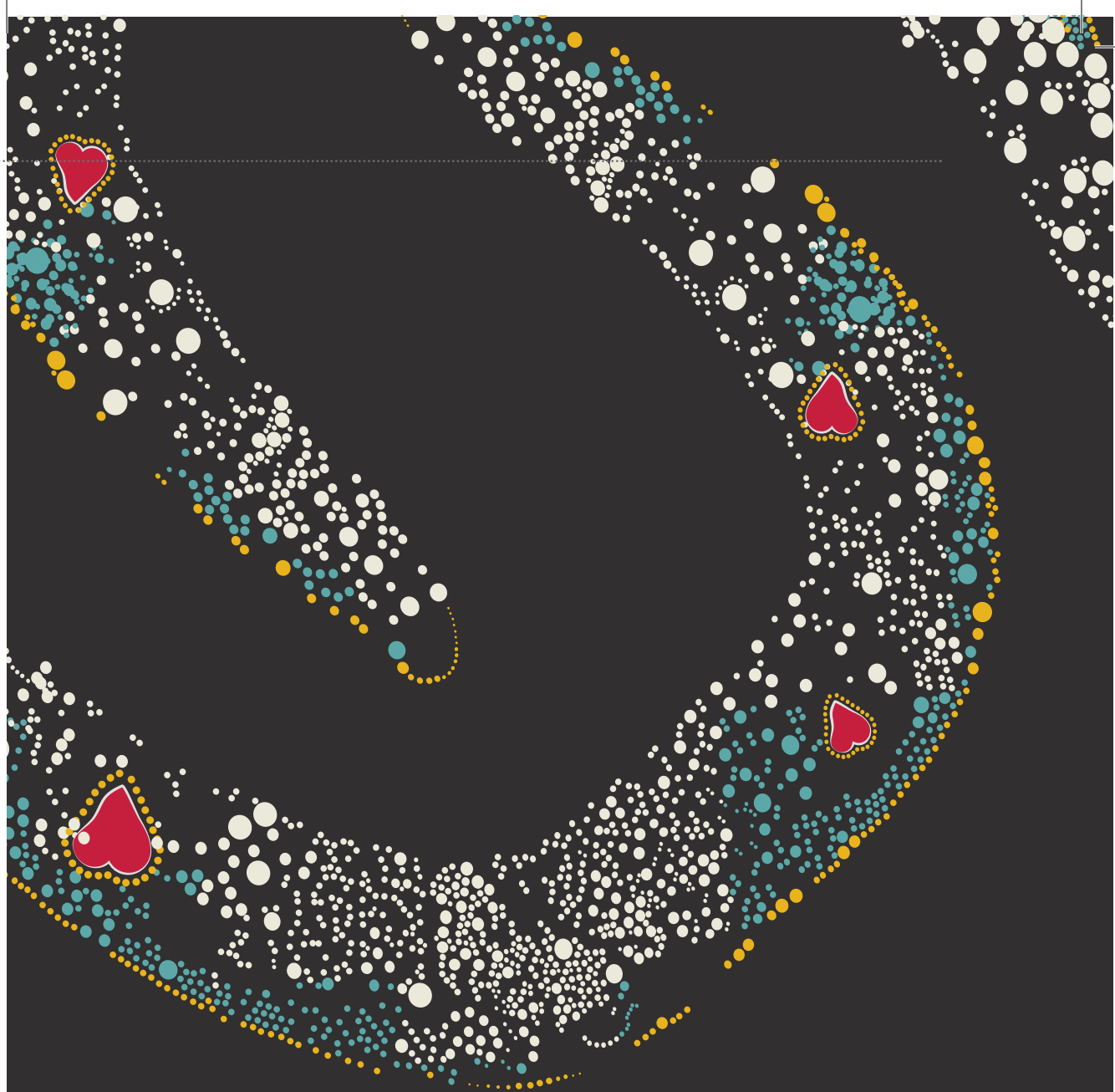
Outlook

In this thesis, we have demonstrated that aggregation-prone proteins influence ncRNA homeostasis, thereby introducing the concept that aggregation-prone proteins have effects that go beyond protein homeostasis. This demonstrates that it may be worth examining how the non-coding genome is altered in neurodegenerative diseases. We have also identified a new aggregation-promoting factor and hope to have uncovered yet another mechanism exemplifying how protein homeostasis can go awry during pathogenesis. Finally, we have provided a starting point for translational research, in which we intend to explore the possibility of a genetic modifier – discovered in *C. elegans* and conserved throughout evolution – being a potential therapeutic target for neurodegenerative diseases.

References

1. Yu Y, Chi B, Xia W, et al. (2015) U1 snRNP is mislocalized in ALS patient fibroblasts bearing NLS mutations in FUS and is required for motor neuron outgrowth in zebrafish. *Nucleic Acids Research* 43:3208–3218. doi: 10.1093/nar/gkv157
2. Gerbino V, Carri MT, Cozzolino M, Achsel T (2013) Mislocalised FUS mutants stall spliceosomal snRNPs in the cytoplasm. *Neurobiology of Disease* 55:120–128. doi: 10.1016/j.nbd.2013.03.003
3. Bading H (2013) Nuclear calcium signalling in the regulation of brain function. *Nat Rev Neurosci* 14:593–608. doi: 10.1038/nrn3531
4. Dufourcq P, Chanal P, Vicaire S, et al. (1999) Iir-2, Iir-1 and Iir-26 encode a new class of zinc-finger proteins and are organized in two overlapping operons both in *Caenorhabditis elegans* and in *Caenorhabditis briggsae*. *Genetics* 152:221–235.
5. Schramm L, Hernandez N (2002) Recruitment of RNA polymerase III to its target promoters. *Genes & Development* 16:2593–2620. doi: 10.1101/gad.1018902
6. Clemens KR, Liao X, Wolf V, et al. (1992) Definition of the binding sites of individual zinc fingers in the transcription factor IIIA-5S RNA gene complex. *Proceedings of the National Academy of Sciences* 89:10822–10826.
7. Clemens KR, Wolf V, McBryant SJ, et al. (1993) Molecular basis for specific recognition of both RNA and DNA by a zinc finger protein. *Science* 260:530–533.
8. Theunissen O, Rudt F, Guddat U, et al. (1992) RNA and DNA binding zinc fingers in *Xenopus* TFIIA. *Cell* 71:679–690.
9. Pelham HR, Brown DD (1980) A specific transcription factor that can bind either the 5S RNA gene or 5S RNA. *Proceedings of the National Academy of Sciences* 77:4170–4174.
10. Lin MT, Beal MF (2006) Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. *Nature* 443:787–795. doi: 10.1038/nature05292
11. Firdaus WJJ, Wyttenbach A, Giuliano P, et al. (2006) Huntingtin inclusion bodies are iron-dependent centers of oxidative events. *FEBS Journal* 273:5428–5441. doi: 10.1111/j.1742-4658.2006.05537.x
12. Wyttenbach A, Sauvageot O, Carmichael J, et al. (2002) Heat shock protein 27 prevents cellular polyglutamine toxicity and suppresses the increase of reactive oxygen species caused by huntingtin. *Human Molecular Genetics* 11:1137–1151.
13. Sin O, Nollen EAA (2015) Regulation of protein homeostasis in neurodegenerative diseases: the role of coding and non-coding genes. *Cell Mol Life Sci*. doi: 10.1007/s00018-015-1985-0
14. Ray A, Zhang S, Rentas C, et al. (2014) RTCB-1 Mediates Neuroprotection via XBP-1 mRNA Splicing in the Unfolded Protein Response Pathway. *J Neurosci* 34:16076–16085. doi: 10.1523/JNEUROSCI.1945-14.2014
15. Ballard C, Gauthier S, Corbett A, et al. (2011) Alzheimer's disease. *Lancet* 377:1019–1031. doi: 10.1016/S0140-6736(10)61349-9
16. Takashima A (2009) Amyloid-beta, tau, and dementia. *J Alzheimers Dis* 17:729–736. doi: 10.3233/JAD-2009-1090
17. van Ham TJ, Holmberg MA, van der Goot AT, et al. (2010) Identification of MOAG-4/SERF as a regulator of age-related proteotoxicity. *Cell* 142:601–612. doi: 10.1016/j.cell.2010.07.020
18. Falsone SF, Meyer NH, Schrank E, et al. (2012) SERF protein is a direct modifier of amyloid fiber assembly. *Cell Rep* 2:358–371. doi: 10.1016/j.celrep.2012.06.012
19. Link CD (1995) Expression of human beta-amyloid peptide in transgenic *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences* 92:9368–9372.
20. Radde R, Bolmont T, Kaeser SA, et al. (2006) Abeta42-driven cerebral amyloidosis in

- transgenic mice reveals early and robust pathology. *EMBO Rep* 7:940–946. doi: 10.1038/sj.embor.7400784
21. Serneels L, Van Biervliet J, Craessaerts K, et al. (2009) gamma-Secretase heterogeneity in the Aph1 subunit: relevance for Alzheimer's disease. *Science* 324:639–642. doi: 10.1126/science.1171176
 22. Jankowsky JL, Fadale DJ, Anderson J, et al. (2004) Mutant presenilins specifically elevate the levels of the 42 residue beta-amyloid peptide in vivo: evidence for augmentation of a 42-specific gamma secretase. *Human Molecular Genetics* 13:159–170. doi: 10.1093/hmg/ddh019
 23. Garcia-Alloza M, Robbins EM, Zhang-Nunes SX, et al. (2006) Characterization of amyloid deposition in the APPswe/PS1dE9 mouse model of Alzheimer disease. *Neurobiology of Disease* 24:516–524. doi: 10.1016/j.nbd.2006.08.017
 24. Lalonde R, Kim HD, Maxwell JA, Fukuchi K (2005) Exploratory activity and spatial learning in 12-month-old APP(695)SWE/co+PS1/DeltaE9 mice with amyloid plaques. *Neurosci Lett* 390:87–92. doi: 10.1016/j.neulet.2005.08.028



Summary

Aging-associated neurodegenerative diseases – including Alzheimer's, Parkinson's and polyglutamine diseases – are projected to be among the top four causes of burden of disease by 2030 (World Health Organization, 2006). This demands investment in biomedical research to find disease-modifying treatments and to improve quality of life in the long term.

One of the hallmarks of neurodegenerative diseases is the presence of protein aggregates in the brains of affected patients. At the cellular and molecular level, protein aggregation results from the disruption of protein homeostasis, which is essential to maintain the correct synthesis, folding, shuttling and clearance of proteins in the cell. One major hurdle in the field is that we do not know why protein aggregation occurs and how it correlates to pathogenesis. What we do know is that specific aggregation-prone proteins have been identified for the different neurodegenerative diseases, and genetic studies have enabled the identification of key mutations in their corresponding genes. Currently, efforts are being made at the genomic, transcriptomic and proteomic level to understand what drives protein aggregation and how it is mechanistically linked to disease.

In this thesis, my motivation was to understand what drives protein aggregation and toxicity. Specifically, my aim was to discover what cellular mechanisms occur as a reaction to the presence of an aggregation-prone protein in the cell. In **Chapter II**, we provide a comprehensive overview of different cellular strategies used to cope with aggregation-prone proteins. We show how small model organisms – including yeast, nematodes, and flies – can be valuable tools to identify the genes involved in protein aggregation and toxicity. Importantly, we show that many of these genes are conserved in mice, allowing more complex studies that can be potentially extrapolated to humans. We then introduce an emerging concept – namely the contribution of the non-coding genome to neurodegeneration – and describe the progress made in this field towards a better understanding of the mechanisms of disease. We demonstrate how recently published work has revealed the link between RNA metabolism and neurological and neurodegenerative diseases, and describe how microRNAs, tRNAs and other types of non-coding RNA are dysregulated in these diseases.

One unresolved issue is the lack of a mechanistic explanation for how impaired RNA metabolism determines pathogenesis.

In **Chapter III**, we highlight the model organism *C. elegans* as a versatile tool for gene function studies. In particular, we focus on “humanized” models of *C. elegans* for neurodegenerative diseases – including Alzheimer’s, Parkinson’s and polyglutamine diseases – and describe how they can be used in forward genetic screens (with ethyl methane sulfate) and in reverse genetics (genome-wide RNAi screens). One aspect of these models worth mentioning is that protein aggregation can be uncoupled from protein toxicity. Next, we present a selection of genetic screens performed in *C. elegans* and show that the genetic modifiers of proteotoxicity identified in these screens are often involved in general cellular processes (e.g. protein quality control system, RNA metabolism, cell cycle), demonstrating that protein aggregation and toxicity can result from different levels of cellular dysregulation.

In line with this, **Chapter IV** describes our work that stemmed from a genetic screen performed in a *C. elegans* model for polyglutamine aggregation aimed at finding genes that promoted protein aggregation and toxicity. This led to the discovery of the novel genetic modifier of aggregation *moag-2* (modifier of aggregation-2), a gene that when mutated suppresses protein aggregation up to 51%. Whole-genome sequencing revealed the causative gene of *moag-2* to be *lir-3*. We found that *moag-2/lir-3* encodes a protein with a predicted nuclear localization signal and two non-canonical zinc finger domains, which are homologous to those in the transcription factor for RNA Polymerase III A (TFIIIA). Indeed, our data revealed that MOAG-2/LIR-3 binds to the same genomic regions as does the RNA Polymerase III machinery in its regulation of the transcription of small non-coding RNAs, including small nuclear RNAs, small nucleolar RNAs and transfer RNAs. These small non-coding RNAs are largely involved in the regulation of gene expression or in the modification of other RNA molecules. Once we had determined the function of MOAG-2/LIR-3, we were surprised to learn that MOAG-2/LIR-3 promotes polyglutamine aggregation in a manner that is independent of its function as a transcriptional regulator. We thereby propose a scenario where aggregation-prone proteins can trigger a switch of function in cellular proteins and hijack

them to promote aggregation. Another unprecedented finding from this work is that the presence of polyglutamine expansion proteins induces the downregulation of snRNA, snoRNA and tRNA genes in wild type animals. Although we do not yet know the biological consequences of this observation, it suggests that aggregation-prone proteins can also affect the homeostasis of non-coding RNA.

In **Chapter V**, we explore the transcriptional changes that occur in the cell when it is exposed to aggregation-prone proteins. By performing whole-transcriptome profiling in a *C. elegans* model of polyglutamine diseases, we found that the expression of aggregation-prone proteins delays animal development and, at the same time, triggers genes involved in cellular stress responses.

One important aspect of identifying modifiers of aggregation in small model organisms is their subsequent validation in human cells and mouse models. Previous work from our group identified MOAG-4/SERF as a positive regulator of proteotoxicity in *C. elegans* models of polyglutamine disease, Parkinson's and Alzheimer's disease. In **Chapter VI**, we describe the generation of a knockout mouse for SERF2, one of the human orthologs of MOAG-4. We found that eliminating *Serf2* is not favorable for mouse development, as it frequently results in embryonic lethality. We then explain how this finding led us to generate a brain-specific *Serf2* knockout mouse instead, which yielded viable and fertile animals. We further validate the potential use of these knockout mice by demonstrating that SERF2 was successfully eliminated from the brain without affecting its expression in other organs. Finally, we propose that these brain-specific *Serf2* knockout mice be used to further study the role of SERF2 as a modifier of aggregation in the brain by mating it with mouse models of neurodegenerative diseases (**Chapter VII**).

In this thesis, we have identified a new genetic regulator of protein aggregation. The findings described here will improve not only the exploration of new mechanisms but also our understanding of the complex protein aggregation process. This work has also provided a starting point for transposing findings from small model organisms to mammalian systems, paving the way for translational research in the long term.



Samenvatting

Volgens voorspellingen van de World Health Organization dringen neurodegeneratieve ziekten, zoals de ziekten van Alzheimer en Parkinson, in 2030 door tot de top vier van de aandoeningen die de meeste ziektelast veroorzaken (World Health Organization, 2006). Biomedisch onderzoek naar aangrijpingspunten voor therapie en naar mogelijkheden om de kwaliteit van leven te verbeteren is daarom van groot belang.

Een van de kenmerken van neurodegeneratieve ziekten is de aanwezigheid van eiwitaggregaten in de hersenen van getroffen patiënten. Op cellulair en moleculair niveau wordt eiwitaggregatie veroorzaakt door een verstoring van de eiwithomeostase in de cel, normaal een samenspel van controle op de synthese, de vouwing, het transport en de afbraak van eiwitten in de cel. Een belangrijk probleem in het veld is dat we niet weten wat eiwitaggregatie aandrijft of hoe het verband houdt met de pathogenese. Wat we wel weten is dat er bij de verschillende neurodegeneratieve ziekten, specifieke aggregatie-gevoelige eiwitten een rol spelen en dat mutaties in hun overeenkomstige genen de oorzaak zijn van erfelijke varianten van deze ziekten. Momenteel wordt veel genetisch, celbiologisch en biochemisch onderzoek gedaan om te achterhalen wat de oorzaak is van de toxiciteit en aggregatie van deze ziekte-eiwitten.

Het doel van mijn proefschrift was om te begrijpen welke biologische processen eiwitaggregatie en toxiciteit aandrijven en welke cellulaire mechanismen optreden als reactie op de aanwezigheid van aggregerende eiwitten in de cel. In **Hoofdstuk II**, geven we een uitgebreid overzicht van de verschillende cellulaire strategieën die gebruikt worden om aggregatie-gevoelige eiwitten aan te pakken. We laten zien hoe kleine modelorganismen – zoals gist, nematoden, en vliegen – waardevol kunnen zijn om genen te ontdekken die betrokken zijn bij eiwitaggregatie en de toxiciteit. We laten zien dat veel van deze genen evolutionair geconserveerd zijn in zoogdieren, wat betekent dat de vindingen potentieel kunnen worden geëxtrapoleerd naar mensen. Vervolgens introduceren we de bijdrage van het niet-coderende genoom aan neurodegeneratie en beschrijven recente inzichten op dit gebied. Wij beschrijven de relatie tussen RNA metabolisme en neurologische en neurodegeneratieve ziekten, en geven een overzicht van de betrokkenheid

van microRNAs, tRNA's en andere vormen van niet-coderende RNA bij deze ziekten. Er is nog geen mechanistische verklaring voor hoe een verstoord RNA metabolisme bijdraagt aan de pathogenese.

In **Hoofdstuk III**, hebben we aandacht voor het modelorganisme *C. elegans* als een veelzijdig diervorm om de functie van genen te ontdekken. In het bijzonder richten we ons op "gehumaniseerde" *C. elegans* modellen voor neurodegeneratieve ziekten en hoe ze kunnen worden gebruikt in genetische screens. Interessant is dat in deze modellen eiwitaggregatie en toxiciteit losgekoppeld kunnen zijn. Vervolgens presenteren we een selectie van genetische screens die zijn uitgevoerd in *C. elegans* en die hebben aangetoond dat de genetische regulatoren van eiwitaggregatie en toxiciteit vaak betrokken zijn bij algemene cellulaire processen (bijvoorbeeld eiwithomeostase, RNA stofwisseling, celcyclus), waaruit blijkt dat eiwitaggregatie en toxiciteit kunnen voortvloeien uit verschillende niveaus van cellulaire ontregeling.

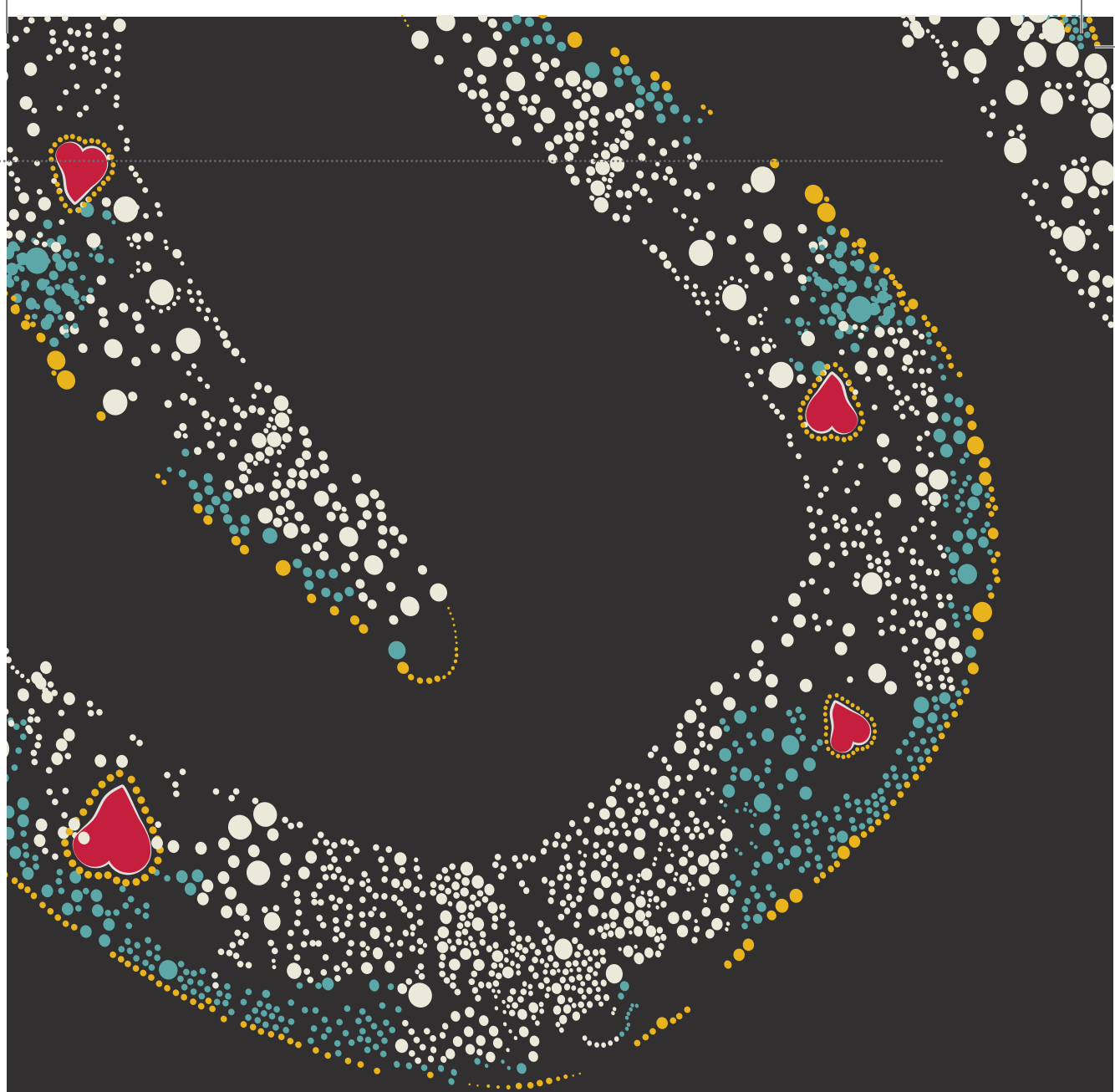
Hoofdstuk IV beschrijft een vervolg op een genetische screen in een *C. elegans* model voor polyglutamine aggregatie: de identificatie van de genetische mutant *moag-2* (modifier of aggregation-2) die eiwitaggregatie voor meer dan 50% vermindert. Met behulp van genoom sequencing hebben we gevonden dat de mutatie in *moag-2* ligt in het gen *lir-3*, waarvan de functie onbekend is. MOAG-2/LIR-3 lijkt te coderen voor een eiwit met een nucleair lokalisatiesignaal en twee zinkvinger domeinen die verwant zijn aan de zinkvinger domeinen in de transcriptiefactor voor RNA Polymerase IIIA (TFIIIA). Met behulp van een eiwit-DNA interactiestudie, vonden we dat MOAG-2/LIR-3 bindt aan dezelfde DNA sequenties als RNA Polymerase III en ook de transcriptie reguleert van kleine niet-coderende RNA's, waaronder kleine nucleaire RNA's (snRNAs), kleine nucleolaire RNAs (snoRNAs) en transfer RNAs (tRNAs). Deze kleine RNA's zijn betrokken bij de regulatie van genexpressie of bij de bouw van andere RNA-moleculen. We waren verrast dat MOAG-2/LIR-3 polyglutamine aggregatie aanstuurt op een manier die onafhankelijk is van zijn functie als regulator van transcriptie. Aggregatie-gevoelige ziekte-eiwitten lijken dus op twee manieren een schadelijk effect te kunnen hebben: ze inactiveren belangrijke cellulaire eiwitten en gebruiken deze om hun eigen aggregatie te versterken. Een andere verrassende ontdekking was dat de aanwezigheid van polyglutamine de

hoeveelheid snRNA, snoRNA en tRNA reduceerde in de worm. Hoewel we nog niet weten wat de biologische gevolgen hiervan zijn en of dit ook te maken heeft met *moag-2/lir-3*, suggereren de resultaten dat aggregatie-gevoelige eiwitten invloed hebben op de homeostase van niet-coderende RNAs.

In **Hoofdstuk V** verkennen we de transcriptionele veranderingen die optreden in *C. elegans* wanneer deze aggregerende eiwitten aanmaken. Met behulp van een transcriptoom analyse in een *C. elegans* model van polyglutamine ziekte, vonden we een effect op de expressie van genen die betrokken zijn bij groei en ontwikkeling en op genen die betrokken zijn bij cellulaire reactie op stress.

Om te bepalen of modifiers van aggregatie in kleine modelorganismen ook een rol spelen bij ziekten, is het belangrijk om eerst hun rol in menselijke cellen en zoogdiermodellen te onderzoeken. Eerder identificeerde we MOAG-4/SERF als positieve regulator van eiwitaggregatie en toxiciteit in *C. elegans* modellen voor polyglutamine, Parkinson en Alzheimer. **Hoofdstuk VI** beschrijft hoe we een knockout muis hebben gemaakt voor SERF2, één van de humane orthologen van MOAG-4. We vonden dat het uitschakelen van SERF2 niet gunstig is voor ontwikkeling van de muis, omdat het meestal leidt tot embryonale letaliteit. We leggen uit hoe we vervolgens een brein-specifieke SERF2 knockout muis hebben gemaakt die wel levensvatbaar en vruchtbaar is. We laten zien dat we SERF2 met succes uit de hersenen hebben gehaald, zonder dat de expressie in andere organen is aangetast. Ten slotte stellen wij voor dat deze hersenspecifieke SERF2 knockout muizen kunnen worden gebruikt voor het verder bestuderen van de rol van SERF2 als modifier van aggregatie in de hersenen door ze genetisch te combineren met muismodellen voor neurodegeneratieve ziekten (**Hoofdstuk VII**).

Met de identificatie van een nieuwe genetische regulator van eiwitaggregatie opent dit proefschrift mogelijkheden voor onderzoek naar mechanismen die eiwitaggregatie en toxiciteit in cellen kunnen verklaren. Bovendien biedt het met de generatie van een muismodel een uitgangspunt voor het vertalen van bevindingen in *C. elegans* naar zoogdierenmodellen voor neurodegeneratieve ziekten, om in de toekomst toe te werken naar translationeel onderzoek.



Resumo

A doença de Alzheimer, a doença de Parkinson e as doenças de expansão de poliglutaminas são doenças neurodegenerativas que se prevê virem a ser a quarta principal causa de custos de saúde em 2030 (World Health Organization, 2006). É, portanto, necessário investir na investigação biomédica destas patologias de forma a encontrar tratamentos que permitam melhorar a qualidade de vida a longo prazo.

Uma característica diagnosticante das doenças neurodegenerativas é a presença de agregados proteicos no cérebro. Estas estruturas resultam de uma incapacidade, a nível celular e molecular, da manutenção da homeostase proteica, que decorre por sua vez de disfunções em vias envolvidas na síntese, *folding*, tráfego e degradação de todas as proteínas celulares. Ainda não se conhece a razão por que ocorre a agregação proteica observada, e de que modo é que esta última está relacionada com a patologia. São, no entanto, conhecidas, para determinadas doenças neurodegenerativas, as proteínas que têm predisposição para agregar (as chamadas *aggregation-prone proteins*) e, através de estudos genéticos, já foram identificadas mutações nos respectivos genes que determinam o desenvolvimento da patologia. Vários estudos realizados a nível genómico, transcriptómico e proteómico estão a ser efetuados no sentido de compreender os fatores envolvidos na agregação proteica e de que modo é que esta está mecanisticamente relacionada com a patologia.

O objectivo principal desta tese foi o estudo dos mecanismos que promovem a agregação e toxicidade proteica em doenças neurodegenerativas. Mais concretamente, procurou-se compreender como é que a célula reage perante a presença de proteínas com predisposição para agregar. O **Capítulo II** descreve as diferentes estratégias celulares que existem para lidar com este tipo de proteínas. Nesse capítulo também são sublinhadas as vantagens de utilizar modelos animais — nomeadamente leveduras, nemátodes e moscas — como ferramentas versáteis para descobrir genes (ou moduladores genéticos) envolvidos na agregação e toxicidade proteica. Muitos dos genes descobertos nestes organismos estão conservados em ratinhos, o que permite estudos mais complexos e cujas descobertas poderão ser potencialmente aplicáveis a humanos. No **Capítulo II**, explora-se ainda o conceito emergente da contribuição do genoma não-codificante para a neurodegeneração, e descreve-

se o progresso realizado nesta área de investigação. São apresentados vários exemplos recentes onde se demonstra a relação entre o metabolismo de RNA e doenças neurológicas e neurodegenerativas e, mais concretamente, descreve-se como é que disfunções em microRNAs, tRNAs e outros tipo de RNAs se correlacionam com a patologia. A contribuição das disfunções no metabolismo do RNA para o desenvolvimento da patologia e os mecanismos envolvidos são questões que permanecem em aberto.

No **Capítulo III**, apresenta-se o modelo animal *C. elegans* como uma ferramenta para descobrir a função de genes. Explora-se como se pode tirar partido de modelos de *C. elegans* para as doenças de Alzheimer, Parkinson ou Huntington, e descreve-se como podem ser utilizados em *screens* genéticos (utilizando etil metano sulfonato ou RNA de interferência). Um aspeto a mencionar é que, nestes modelos, a agregação e toxicidade proteicas são fenótipos que podem ser analisados independentemente um do outro. Ainda no **Capítulo III** são apresentados resultados de *screens* genéticos nos quais se observa que moduladores genéticos de proteotoxicidade estão frequentemente associados a funções metabólicas gerais (sistema de controlo de qualidade proteica, metabolismo de RNA, ciclo celular, etc.), demonstrando que a agregação e toxicidade proteica podem resultar de diferentes níveis de disfunção celular.

O **Capítulo IV** descreve o trabalho desenvolvido a partir de um *screen* genético num modelo de *C. elegans* para doenças de expansão de poliglutaminas. Neste *screen* procuraram-se genes que promovem a agregação e toxicidade proteica, usando a poliglutamina como exemplo de uma proteína com predisposição para agregar. Neste capítulo descreve-se também a descoberta do modulador genético *moag-2* (*modifier of aggregation-2*) que, quando mutado, reduz a agregação proteica até 51%. Após sequenciação, descobriu-se que o modulador genético *moag-2* corresponde ao gene *lir-3*. O gene *moag-2/lir-3* codifica uma proteína de função desconhecida, mas a sua sequência de aminoácidos contém um sinal de localização nuclear e dois domínios com “dedos” de zinco (*zinc fingers*) homólogos aos do factor de transcrição para a RNA Polímerase III A (TFIIIA). Neste estudo, descobriu-se ainda que a proteína MOAG-2/LIR-3 se liga aos mesmos locais genómicos

que a maquinaria da RNA Polímerase III para transcrição de pequenos RNAs não-codificantes — entre os quais se incluem os *small nuclear RNA* (snRNA); *small nucleolar RNAs* (snoRNAs) e RNAs de transferência (tRNA). A função destes pequenos RNAs não-codificantes é a regulação da expressão genética ou modificação de outras moléculas de RNA. Curiosamente, descobriu-se que a proteína MOAG-2/LIR-3 promove a agregação de poliglutaminas de uma forma que é independente da sua função como factor de transcrição. Assim, proteínas com predisposição para agregar conseguirão sequestrar e converter proteínas celulares normais em factores que promovem agregação proteica. Uma descoberta adicional deste trabalho foi que proteínas com predisposição para agregar reduzem a abundância de pequenos RNA não-codificantes em animais *wild type*, nomeadamente snRNAs, snoRNAs e tRNAs. Apesar de serem desconhecidas as consequências biológicas desta observação, os resultados sugerem que proteínas com predisposição para agregar também podem influenciar a homeostase de RNAs não-codificantes.

O **Capítulo V** descreve as alterações transcricionais que ocorrem quando células expressam proteínas com predisposição para agregação proteica. Ao comparar o perfil transcriptómico de animais *wild type* com animais que expressam poliglutamina, descobriu-se que esta atrasa o desenvolvimento animal — consistentemente com observações anteriores de que estes últimos demoram normalmente mais tempo a atingir o mesmo estágio de desenvolvimento do que animais *wild type*. Resultados obtidos pelo mesmo estudo revelaram ainda quais são as vias de sinalização activadas como consequência da presença de poliglutaminas. Mostra-se também que a expressão de proteínas com predisposição para agregar induz a activação da *unfolded protein response* (UPR), uma das principais vias envolvidas na manutenção da homeostase proteica.

Uma vez identificados moduladores em modelos animais como leveduras, nemátodes ou moscas, é necessário validar a sua função em mamíferos ou cultura de células humanas. A proteína MOAG-4/SERF foi identificada, em trabalho previamente publicado pelo nosso grupo, como um modulador genético de agregação proteica em modelos de *C. elegans* para a doença de Alzheimer, de Parkinson e de expansão de poliglutaminas. O **Capítulo VI**,

descreve a geração de um modelo de murganho com deleção (*knockout*) do gene SERF2, um dos homólogos de MOAG-4. Descobriu-se que a deleção de SERF2 não é favorável para o desenvolvimento animal, uma vez que resultou frequentemente em letalidade embrionária com penetrância incompleta. Esta observação levou a que se tentassem gerar mutantes condicionais nos quais a expressão de SERF2 é eliminada apenas no cérebro. Esta estratégia foi bem sucedida uma vez que se obtiveram mutantes condicionais viáveis e férteis. Mostra-se ainda que a deleção de SERF2 do cérebro foi conseguida sem afectar a sua expressão noutros órgãos. Finalmente, propomos que estes mutantes condicionais possam ser cruzados com modelos de murganho para a doença de Alzheimer, o que permitirá estudar a função de SERF2 como modulador de agregação proteica na doença de Alzheimer (**Capítulo VII**).

Em suma, esta tese descreve a descoberta e caracterização de um novo modulador genético para a agregação proteica e propõe mecanismos alternativos que podem ajudar a compreender melhor a complexidade do processo de agregação proteica. Paralelamente, o trabalho aqui apresentado deu início ao processo de validação de um modulador genético descoberto em *C. elegans* e a sua transposição para modelos de Alzheimer em murganhos, abrindo caminho a investigação translacional.



Acknowledgments

I want to first and foremost show my gratitude to the **GABBA PhD Program** of the University of Porto. I will never forget my interview with **Prof. Maria de Sousa, Prof. Graça Porto, Prof. Fátima Carneiro, Prof. Alexandre do Carmo, Prof. Sobrinho Simões** and **Prof. João Relvas**, which started off exchanging impressions about God and ended somehow with a group photo. I am forever grateful for being part of the GABBA family and for all the opportunities that came with that. A special “thank you” to Prof. Alexandre do Carmo and Prof. António Amorim for always facilitating the great bureaucratic challenge that was my co-tutorship between the University of Porto and the University of Groningen. **João**, thank you for being my co-supervisor at Porto.

Dear **Ellen**, I owe you my deepest and heartfelt gratitude. Thank you for taking a chance on me and for guiding and shaping me into who I am today. I admire your generosity and trust towards people, and how you propel them to aim higher and achieve better. I am grateful for your patience and constant optimism whenever I felt too impatient or doubtful about my own projects. You opened so many doors for me and introduced me to the joy of teaching. With you, I had the space and freedom to do science without restraints, and I am so happy for that. Doing science can question one’s perseverance and self-confidence, but the truth is that every high and every low in this adventure was worth it, and I leave even more passionate about what I do. Thank you for trusting me and I wish you the best of luck and success for the exciting stories that will come out of the lab.

“Heeeeeeey” **Victor**, your expertise in “omics” was pivotal to the MOAG-2/LIR-3 story and I am so happy you were around to help, in your calm, cool way of being. Many of my favorite work discussions were with you and I appreciate how you always listened and valued my ideas and thoughts. Remember when we hi-fived after we linked MOAG-2/LIR-3 to Box A and Box B? That moment will always stay close to my heart.

I am thankful to the reading committee, **Prof. dr. Ody Sibon, Prof. dr. Gerald de Haan** and **Prof. dr. Sander van Heuvel** for their time and effort to critically evaluate this thesis. I hope you enjoyed reading it as much as I did writing it.

I would like to show my gratitude to **Michelle Kudron**. The OP312 strain and the ChIP seq data on MOAG-2/LIR-3 were fundamental for the success of the project and

I appreciate your willingness to share these with us without any reservations from the start.

Dear **Kohta Ikegami**, reading your Molecular Cell paper (2013) was one of my first “aha” moments while working on the MOAG-2/LIR-3 story. It shaped a lot of what happened afterwards and it was wonderful to count with your help along the way. I wish you a very successful career and I hope we may continue to collaborate in the future.

When I joined the lab, I was supposed to work on a mouse project for my PhD. **Bart van de Sluis**, thank you for sharing your expertise and guidance at the very beginning, which helped us take off with this project. I will always imagine you jumping off a helicopter and slalom like crazy in a snowy mountain... with your kids following you.

I would like to thank **Juul Baller** at the CDP for mediating all matters concerning the mouse project. I hope to welcome you someday in Portugal and show what a beautiful country it is!

I would like to thank the Genetics department for the welcoming and friendly atmosphere, it was great to share the office and lab with you during the first two years of my PhD. A special “thank you” to **Jan Osinga** (for help with the sequencing and for practicing my Dutch), **Bahram Sanjabi** (for constant kindness), **Michiel Fokkens** (for being funny all the time) and my favorite officemate **Duco Schriemer**, with whom I shared extensive thoughts of football tactics (cof cof... still need to work on that powerpoint where Portugal beats the Netherlands).

It was also at the Genetics department where I met people who brought so much fun and laughter to my life in Groningen. **Barbara**, thank you for your friendship during my first two years in Groningen and, importantly, for your support when I needed to take care of matters back home. **Javier**, mamacito, I will always remember your cool way of being and how nice you were to me every time. **Juha**, though we take opposite perspectives in everything, we could still find each other in common ground. I had the best time ice-skating with you. **Cleo**, you are fire and sweetness at the same time, spiced with irony in your sense of humor. The result of this recipe? The awesome creation of Captain Rainbow. Thanks for all the fun. **Luz**, you are such a kind person,

it was always so pleasant to talk to you and share your company. I am glad you can continue your career back home and I hope life always smiles at you. **Mamen**, you are such a joyful person, you bring light and laughter everywhere you go. It was so nice to have someone like you around, if there were more people like you the world would be a happier place.

Rodrigo, Rodrigo... You were – by far – the funniest soul I've encountered during my life in Groningen and so many of the best moments were spent with you. I keep fond memories of our road trip with Ale, Juha and photobombing with Jimbo across the Netherlands (I will never forgive you for purposely losing Jimbo). How about Rock Werchter, where you discovered how great it was to have noodles... every day? Even to this day, on the other side of the ocean, you make me laugh like nobody else. I imagine a rainbow shining wherever you go, with sprinkles of colorful confetti falling upon you. There will never be rain, only chocolate thunder.

Anna Duarri, I started speaking Portuguese with you! We had so much fun together back in the day, fooling around. I miss our moments at the Pacific, while drinking our favorite cocktail (maybe two, maybe three, maybe four...) and dancing the night away at Enzo. You were also great travelling companion and I hope to welcome you someday in Portugal so we can go for sushi!

My dear **Isis**. You are a beautiful human being with a very big heart. I am so happy that we became friends and that you could share a little bit of you with me. Our friendship has lasted throughout these years and I want to thank you for being there for me in the crazy and happy moments, chiquitina. You are sunshine! I look forward to welcome you to my home in Portugal so we can enjoy the warmth of the sun by the sea.

I am grateful to the past, current and new members of the lab. I believe that being in this lab has allowed me to mature as a scientist and as a person, and I have learned from each and every one of my colleagues. I am grateful for their kindness and patience during these five years. First, I would like to thank the people who welcomed me when I joined the lab: **Karen, Annemieke, Mats, Céline** and **Eva**.

Renéetje! Pragmatic, organized and efficient is what defines you and I was so lucky to count with your abilities in the lab. Thank you so much for your time, dedication and effort. I sincerely hope that you had fun doing experiments by my side as I did with you. And I also hope that I packed you with enough random trivia knowledge (oh yes) to kickass at pub quizzes. Schaatje, I wish you the best that life can offer and I look forward to your first aquarium in your new home!

Esthertje! I like your sense of humor and I am very happy that our friendship extended beyond the lab. I keep fond memories of Wim's cups, game nights, cupcakes, movie nights, sushi dinners, dancing at the Warhol... thank you for so many moments well spent. I promise that if I ever get a smartphone, you will be one of the first to get a whassup/what's up/whatsapp message. I wish you the very best in your career and I hope we find each other again in the future.

Wytse, our paths crossed briefly with the mouse project and I'm happy to have had you alongside setting up this challenging project. It was ok, right? I will miss exchanging impressions about movies and music with you! **Kees**, though we only met for brief months, I appreciated your happy mood and homemade cakes. I wish you success in your future endeavors!

Helen, I will miss your pumpkin soup most definitely your rhubarb jam! Plus, you were only one with whom I could share the joys of having dogs and cats! I wish you success finishing your PhD and future endeavours.

I would also like to thank my former student – **Hai Hui** – for his initiative, dedication and important contribution to the MOAG-3/LIR-3 story. It was a pleasure working with you and your independence, discipline and initiative were exemplary.

To the new generation in the lab – dear **Anita**, super **Mandy** and freaky **Freek** – it is wonderful to see your enthusiasm, dedication and good sense of humor at work! You bring positive energy to the lab, with lots of smiles and laughter. Make the most of your talent and I wish you exciting discoveries during your PhD!

After moving the lab to ERIBA, I met some people I would like to thank for their kindness and friendship. **Marianna Bevova**, we first met six years ago while you were giving a course on human genetics to the GABBA students in Porto. What a wonderful surprise it was to cross paths with you again in Groningen. Best yet, I had the pleasure of lecturing the new GABBA students with you and Yuri. I look up to you and to hear that you believe in me by propelling me to continue doing research means a lot to me (that also means at least one person thinks I'm minimally capable to do it, haha). **Diana Spierings**, we seemed to hit it off right from the beginning of our friendship. You were always so nice to me and I really liked it whenever you came by to make me company while I was writing my thesis. **Yin Fai**, or shall I say, Yin-fire? Thank you for your friendship and for sharing your good mood with me. You are truly dedicated to your work, organized and precise in everything you do and I respect how you stay afterhours many times to help the research of others become a success. Thank you for all the sushi boxes and especially for your wonderful, delicious *ho-fan*. It is my sincere wish that you have good fortune and... a stable future, of course! I hope to welcome you one day to Portugal.

Tristan, it was real fun to work with you during the final stages of my PhD, and I am very grateful for your help with the RNA seq data. Always available, always helpful, anytime, anywhere. If I could, I would give you your weight in chocolates and candy. Looking forward for a Star Wars meets X-Men movie, I have my doubts whether a light saber can really cut Wolverine's adamantium claws. I wish you all the best for your PhD. **Roméo**, I really enjoyed working with you and I thank you for your dedication and enthusiasm helping out with the ChIP seq data. You are very talented and I know you will go far. I am so happy to be friends with you and I sincerely wish you the best in everything you decide to do.

Amr, Amr, Amr, savah al-hair! I am so very happy to have met you. You are organized, methodic and disciplined in the lab, and your blots always look amazing. Thank you for sharing your culture with me, I absolutely adored our long conversations about our cultures and ways of thinking. I will never forget hearing "good job!" as you walked by my bench, it always made me smile and feel very positive about my experiments. Professor Amr, I wish you a very successful PhD. Good job!

Baobei! **Xin Xuan**, you were one of the funniest things that happened to me. It was so much fun to have you around and live so many *pokai* moments together in the lab. I mean, come on. *Pokai* this, *pokai* that. *Pokai* all the time. You certainly awakened the Asian side of me. It was always so nice when you surprised me with food treats or bubble tea, thank you friend. I think we can say that one of our great gastronomic achievements was Chi Chi's, so we should totally go back for *twee halen een betalen* margaritas, or should I say, super mega large margaritas? **Rik**, my Nigerian friend. Thank you for the super awesome movie nights, *mah jong* and game evenings, and especially for the festive meals at Asian restaurants with four persons occupying three tables of never ending food!

KLAAS! It was such a joy to meet you during the final stages of my PhD! Looking back, I cannot believe how you handled with calm and “zen-ness” this little ball of concentrated energy bouncing up and down in the microscopy center. Thank you for teaching me the basics of confocal microscopy and don't worry, I never made any karate moves in the Leica Sp8.

My thanks to the **GABBA 13** people, with whom I spent the first year of my PhD. I keep fond memories of our year in Porto and – most of all – of our time in rainy Sevilla. You are bright minds with bright futures ahead of you, and this makes me proud of the quality of people we export around the world. A special hug to **Mariana** and **João Dias**.

I am very proud of the design of this thesis, which would not have been possible without the talent of **Clau**. Dearest Clau, I am so happy that our paths crossed and that our friendship has remained throughout these years. I admire your courage, strength and determination. You are a beautiful and amazing person and I admire you very much. I wish you and Simonita the best of futures and all the happiness in the world. I hope that we may meet again someday in the future.

My thanks to **Roger Franke**, for taking care of the printing of this thesis at NetzoDruk. Our friendship goes back to my days in Gopher, and I am very happy that this collaboration has remained. I wish you prosperity in your business and health and happiness for you and your family.

Now to my paranymphs. Dear **Manel**, you were the very first person who took care of me when I arrived to Groningen. I still remember when you picked me up at the train station and showed me to my new house. Always smiling, always in a good mood and always straight to the point with your supersonic speech. Thank you for all the Portuguese meals you cooked, they were delicious and made me miss home a little bit less. I really do think the Dutch way of life suits you and I can only wish for even more successful accomplishments in your career (adding to your Marie Curie and Veni!). How could you expect otherwise, when you come from good families? Hahaha!

Ale, where should I start? You were one of the best things that happened to me. You were my colleague, my neighbor, my friend and – above all – my big brother. You understood me and were always there for me. Thank you for listening to my fears, complaints and expectations, for lifting me up when I felt less happy and for taking care of me. Thank you and sorry for dealing with the Hulk version of me so many times. Thank you for borrowing me your oven and for trying my cupcakes, even with all the complicated decoration. You are bright and always thinking about how things work, and your enthusiasm for science is genuine. I am so glad I could do science by your side because I feel like I learned from a true scientist. I know you have many qualities that go beyond science (awesome *tortilla de patatas* and delicious *salmorejo*) but I hope to see you become a rock star group leader holding a PRS guitar on one hand and a Framus on the other and kickass with cool experiments, many of which you already have in mind but just need the space to do it. Otherwise we can always team up and start up a farm in the Portuguese/Spanish border. Obrigado por TUDO, Ale.

My final and biggest thanks to you, querida mãe. You are such a brave woman and throughout your life you have done and achieved amazing things that go beyond your seemingly tiny figure. You are a fighter and a survivor, and I am so very proud to be your daughter. I always find comfort and guidance in your wisdom, even though it means hearing very often the hard truth. Every moment I felt broken, you elevated my spirit like nobody else. Every time I was too proud to listen, you pushed my ears and reminded me to be a better person. Most of all, thank you for letting me go and explore what was out there without ever holding me back. I am so blessed to have such a wonderful family with you, Kiu Kiu, Kitty and Saikoh. I cherish every single moment with you because you are my absolute happiness. Adoro-te!



List of publications

Sin O, de Jong TV, Kudron M, Seinstra R, Mata-Cabana A, Martineau C, Willinge Prins R, Wang HH, Hogewerf W, Reinke V, Guryev V, Nollen EEA. Functional sequestration of MOAG-2/LIR-3 by polyglutamine expansion proteins. *Manuscript in preparation*

Sin O and Nollen EEA. Regulation of protein homeostasis in neurodegenerative diseases: the role of coding and non-coding genes. *Cellular and Molecular Life Sciences* 2015 Nov; 72(21):4027-47. doi: 10.1007/s00018-015-1985-0. Epub 2015 Jul 21.

Sin O, Michels H, Nollen EEA. Genetic screens in *Caenorhabditis elegans* models for neurodegenerative diseases. *Biochimica et Biophysica Acta* 2014 Oct; 1842(10):1951-1959. doi: 10.1016/j.bbadis.2014.01.015.